A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY8, EXPRESSED HIGHLY IN BRAIN

[0001] This application claims benefit to provisional application U.S. Serial No. 60/248,285, filed November 14, 2000; to provisional application U.S. Serial No. 60/268,581, filed February 14, 2001; to provisional application U.S. Serial No. 60/308,285, filed July 27, 2001; and to provisional application U.S. Serial No. 60/317,166, filed September 4, 2001.

FIELD OF THE INVENTION

[0002] The present invention relates to the fields of pharmacogenomics, diagnostics and patient therapy. More specifically, the present invention relates to methods of diagnosing and/ or treating diseases involving the Human G-Protein Coupled Receptor, HGPRBMY8.

BACKGROUND OF THE INVENTION

[0003] It is well established that many medically significant biological processes are mediated by proteins participating in signal transduction pathways that involve G-proteins and/or second messengers, e.g., cAMP (Lefkowitz, Nature, 351:353-354 (1991)). Herein these proteins are referred to as proteins participating in pathways with G-proteins or PPG proteins. Some examples of these proteins include the GPC receptors, such as those for adrenergic agents and dopamine (Kobilka, B. K., et al., PNAS, 84:46-50 (1987); Kobilka, B. K., et al., Science, 238:650-656 (1987); Bunzow, J. R., et al., Nature, 336:783-787 (1988)), G-proteins themselves, effector proteins, e.g., phospholipase C, adenylate cyclase, and phosphodiesterase, and actuator proteins, e.g., protein kinase A and protein kinase C (Simon, M. I., et al., Science, 252:802-8 (1991)).

[0004] For example, in one form of signal transduction, the effect of hormone binding is activation of an enzyme, adenylate cyclase, inside the cell. Enzyme activation by hormones is dependent on the presence of the nucleotide GTP, and GTP also influences hormone binding. A G-protein connects the hormone receptors to adenylate cyclase. G-protein was shown to exchange GTP for bound GDP when

activated by hormone receptors. The GTP-carrying form then binds to an activated adenylate cyclase. Hydrolysis of GTP to GDP, catalyzed by the G-protein itself, returns the G-protein to its basal, inactive form. Thus, the G-protein serves a dual role, as an intermediate that relays the signal from receptor to effector, and as a clock that controls the duration of the signal.

[0005] The membrane protein gene superfamily of G-protein coupled receptors has been characterized as having seven putative transmembrane domains. The domains are believed to represent transmembrane a-helices connected by extracellular or cytoplasmic loops. G-protein coupled receptors include a wide range of biologically active receptors, such as hormone, viral, growth factor and neuroreceptors.

[0006] G-protein coupled receptors have been characterized as including these seven conserved hydrophobic stretches of about 20 to 30 amino acids, connecting at least eight divergent hydrophilic loops. The G-protein family of coupled receptors includes dopamine receptors, which bind to neuroleptic drugs, used for treating psychotic and neurological disorders. Other examples of members of this family include calcitonin, adrenergic, endothelin, cAMP, adenosine, muscarinic, acetylcholine, serotonin, histamine, thrombin, kinin, follicle stimulating hormone, opsins, endothelial differentiation gene-1 receptor, rhodopsins, odorant, cytomegalovirus receptors, etc.

[0007] Most G-protein coupled receptors have single conserved cysteine residues in each of the first two extracellular loops which form disulfide bonds that are believed to stabilize functional protein structure. The 7 transmembrane regions are designated as TM1, TM2, TM3, TM4, TM5, TM6, and TM7. TM3 has been implicated in signal transduction.

[0008] Phosphorylation and lipidation (palmitylation or farnesylation) of cysteine residues can influence signal transduction of some G-protein coupled receptors. Most G-protein coupled receptors contain potential phosphorylation sites within the third cytoplasmic loop and/or the carboxyl terminus. For several G-protein coupled receptors, such as the β-adrenoreceptor, phosphorylation by protein kinase A and/or specific receptor kinases mediates receptor desensitization.

[0009] For some receptors, the ligand binding sites of G-protein coupled receptors are believed to comprise a hydrophilic socket formed by several G-protein coupled receptors transmembrane domains, which socket is surrounded by hydrophobic residues of the G-protein coupled receptors. The hydrophilic side of each G-protein coupled receptor transmembrane helix is postulated to face inward and form the polar ligand-binding site. TM3 has been implicated in several G-protein coupled receptors as having a ligand-binding site, such as including the TM3 aspartate residue. Additionally, TM5 serines, a TM6 asparagine and TM6 or TM7 phenylalanines or tyrosines are also implicated in ligand binding.

[0010] G-protein coupled receptors can be intracellularly coupled by heterotrimeric G-proteins to various intracellular enzymes, ion channels and transporters (see, Johnson et al., Endoc. Rev., 10:317-331(1989)). Different G-protein β-subunits preferentially stimulate particular effectors to modulate various biological functions in a cell. Phosphorylation of cytoplasmic residues of G-protein coupled receptors have been identified as an important mechanism for the regulation of G-protein coupling of some G-protein coupled receptors. G-protein coupled receptors are found in numerous sites within a mammalian host.

[0011] G-protein coupled receptors (GPCRs) are one of the largest receptor superfamilies known. These receptors are biologically important and malfunction of these receptors results in diseases such as Alzheimer's, Parkinson, diabetes, dwarfism, color blindness, retinal pigmentosa and asthma. GPCRs are also involved in depression, schizophrenia, sleeplessness, hypertension, anxiety, stress, renal failure and in several other cardiovascular, metabolic, neural, oncology and immune disorders (F. Horn and G. Vriend, J. Mol. Med., 76: 464-468 (1998)). They have also been shown to play a role in HIV infection (Y. Feng et al., Science, 272: 872-877 (1996)). The structure of GPCRs consists of seven transmembrane helices that are connected by loops. The N-terminus is always extracellular and C-terminus is intracellular. GPCRs are involved in signal transduction. The signal is received at the extracellular N-terminus side. The signal can be an endogenous ligand, a chemical moiety or light. This signal is then transduced through the membrane to the cytosolic side where a heterotrimeric protein G-protein is activated which in turn elicits a

response (F. Horn et al., Recept. and Chann., 5: 305-314 (1998)). Ligands, agonists and antagonists, for these GPCRs are used for therapeutic purposes.

[0012] The present invention provides a newly discovered G-protein coupled receptor protein, which may be involved in cellular growth properties in brain-related tissues based on its abundance found in the brain for this receptor. The present invention also relates to newly identified polynucleotides, polypeptides encoded by such polynucleotides, the use of such polynucleotides and polypeptides, as well as the production of such polynucleotides and polypeptides. More particularly, the polypeptides of the present invention are human 7-transmembrane receptors. The invention also relates to inhibiting the action of such polypeptides.

SUMMARY OF THE INVENTION

[0013] The present invention provides a novel human member of the G-protein coupled receptor (GPCR) family (HGPRBMY8). Based on sequence homology, the protein HGPRBMY8 is a candidate GPCR. Based on its protein sequence information, the HGPRBMY8 contains seven transmembrane domains, which is a characteristic structural feature of GPCRs. The GPCR of this invention is closely related to the somatostatin and GPR24 receptor families based on sequence similarity using the BLAST program. This orphan GPCR is expressed highly in brain.

[0014] It is an object of the present invention to provide an isolated HGPRBMY8 polynucleotide as depicted in SEQ ID NO:1.

[0015] It is also an object of the present invention to provide the HGPRBMY8 polypeptide, encoded by the polynucleotide of SEQ ID NO:1 (CDS=1 to 1524) and having the amino acid sequence of SEQ ID NO:2, or a functional or biologically active portion thereof.

[0016] It is a further object of the present invention to provide compositions comprising the HGPRBMY8 polynucleotide sequence, or a fragment thereof, or the encoded HGPRBMY8 polypeptide (MW=56.7Kd), or a fragment or portion thereof. Also provided by the present invention are pharmaceutical compositions comprising at least one HGPRBMY8 polypeptide, or a functional portion thereof, wherein the

compositions further comprise a pharmaceutically acceptable carrier, excipient, or diluent.

[0017] It is an object of the present invention to provide a novel, isolated, and substantially purified polynucleotide that encodes the HGPRBMY8 GPCR homologue, or fragment thereof. In a particular aspect, the polynucleotide comprises the nucleotide sequence of SEQ ID NO:1. The present invention also provides a polynucleotide sequence comprising the complement of SEQ ID NO:1, or variants thereof. In addition, the present invention features polynucleotide sequences, which hybridize under conditions of moderate stringency or high stringency to the polynucleotide sequence of SEQ ID NO:1.

[0018] It is an object of the present invention to further provide a nucleic acid sequence encoding the HGPRBMY8 polypeptide and an antisense of the nucleic acid sequence, as well as oligonucleotides, fragments, or portions of the nucleic acid molecule or antisense molecule. Also provided are expression vectors and host cells comprising polynucleotides that encode the HGPRBMY8 polypeptide.

[0019] It is an object of the invention to provide methods for producing a polypeptide comprising the amino acid sequence depicted in SEQ ID NO:2, or a fragment thereof, comprising the steps of a) cultivating a host cell containing an expression vector containing at least a functional fragment of the polynucleotide sequence encoding the HGPRBMY8 protein according to this invention under conditions suitable for the expression of the encoded polypeptide; and b) recovering the polypeptide from the host cell.

[0020] It is also an object of the invention to provide antibodies, and binding fragments thereof, which bind specifically to the HGPRBMY8 polypeptide, or an epitope thereof, for use as therapeutic and diagnostic agents.

[0021] It is a further object of the invention to provide methods for screening for agents which bind to, or modulate HGPRBMY8 polypeptide, e.g., agonists and antagonists, as well as the binding molecules and/ or modulators, e.g., agonists and antagonists, particularly those that are obtained from the screening methods described.

[0022] It is an object of the present invention to also provide a substantially purified antagonist or inhibitor of the polypeptide of SEQ ID NO:2. In this regard,

and by way of example, a purified antibody that binds to a polypeptide comprising the amino acid sequence of SEQ ID NO:2 is provided.

[0023] It is an object of the invention to further provide substantially purified agonists or activators of the polypeptide of SEQ ID NO:2 are further provided.

[0024] It is another object of the present invention to provide HGPRBMY8 nucleic acid sequences, polypeptide, peptides and antibodies for use in the diagnosis and/or screening of disorders or diseases associated with expression of the polynucleotide and its encoded polypeptide as described herein.

[0025] It is a also an object of the present invention to provide kits for screening and diagnosis of disorders associated with aberrant or uncontrolled cellular development and with the expression of the polynucleotide and its encoded polypeptide as described herein.

[0026] It is an object of the present invention to further provide methods for the treatment or prevention of cancers, immune disorders, or neurological disorders involving administering to an individual in need of treatment or prevention an effective amount of a purified antagonist of the HGPRBMY8 polypeptide. Due to its elevated expression in brain, the novel GPCR protein of the present invention is particularly useful in treating or preventing neurological disorders, conditions, or diseases.

[0027] It is an object of the present invention to also provide a method for detecting a polynucleotide that encodes a G-protein coupled receptor, preferably the HGPRBMY8 polypeptide, or homologue, or fragment thereof, in a biological sample comprising the steps of: a) hybridizing the polynucleotide, or complement of the polynucleotide sequence encoding SEQ ID NO:2 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and b) detecting the hybridization complex, wherein the presence of the complex correlates with the presence of a polynucleotide encoding the HGPRBMY8 polypeptide, or fragment therof, in the biological sample. The nucleic acid material may be further amplified by the polymerase chain reaction prior to hybridization.

[0028] It is an object of the instant invention to provide methods and compositions to detect and diagnose alterations in the HGPRBMY8 sequence in tissues and cells as they relate to ligand response.

[0029] It is an object of the present invention to further provide compositions for diagnosing brain-related disorders and for diagnosing or monitoring response to HGPRBMY8 therapy in humans. In accordance with the invention, the compositions detect an alteration of the normal or wild type HGPRBMY8 sequence or its expression product in a patient sample of cells or tissue.

[0030] It is an object of the present invention to provide diagnostic probes for diseases and a patient's response to therapy. The probe sequence comprises the HGPRBMY8 locus polymorphism. The probes can be constructed of nucleic acids or amino acids.

[0031] It is an object of the present invention to further provide antibodies, and immunoreactive portions thereof, that recognize and bind to the HGPRBMY8 protein. Such antibodies can be either polyclonal or monoclonal. Antibodies that bind to the HGPRBMY8 protein can be utilized in a variety of diagnostic and prognostic formats and therapeutic methods.

[0032] It is also an object of the present invention to provide diagnostic kits for the determination of the nucleotide sequence of human HGPRBMY8 alleles. The kits are based on amplification-based assays, nucleic acid probe assays, protein nucleic acid probe assays, antibody assays or any combination thereof.

[0033] It is an object of the instant invention to further provide methods for detecting genetic predisposition, susceptibility and response to therapy related to the brain. In accordance with the invention, the method comprises isolating a human sample, for example, blood or tissue from adults, children, embryos or fetuses, and detecting at least one alteration in the wild type HGPRBMY8 sequence, or its expression product, from the sample, wherein the alterations are indicative of genetic predisposition, susceptibility or altered response to therapy related to the brain.

[0034] It is an additional object of the present invention to provide methods for making determinations as to which drug to administer, dosages, duration of treatment and the like.

[0035] Further objects, features, and advantages of the present invention will be better understood upon a reading of the detailed description of the invention when considered in connection with the accompanying figures/drawings.

BRIEF DESCRIPTION OF THE FIGURES

[0036] Figure 1 shows the full-length nucleotide sequence of cDNA clone HGPRBMY8, a human G-protein coupled receptor (SEQ ID NO:1).

[0037] Figure 2 shows the amino acid sequence (SEQ ID NO:2) from the translation of the full-length HGPRBMY8 cDNA sequence.

[0038] Figure 3 shows the 5' untranslated sequence of the orphan HGPRBMY8 (SEQ ID NO:3).

[0039] Figure 4 shows the 3' untranslated sequence of the orphan HGPRBMY8 (SEQ ID NO:4).

[0040] Figure 5 shows the predicted transmembrane region of the HGPRBMY8 protein where the predicted transmembrane regions, represented by bold-faced and underlined type, correspond to the peaks with scores above 1500.

[0041] Figures 6A- 6J show the multiple sequence alignment of the translated sequence of the orphan G-protein coupled receptor, HGPRBMY8, where the GCG (Genetics Computer Group) pileup program was used to generate the alignment with several known adrenergic and serotonin receptor sequences. The blackened areas represent identical amino acids in more than half of the listed sequences and the grey highlighted areas represent similar amino acids. As shown in Figures 6A- 6J, the sequences are aligned according to their amino acids, where: HGPRBMY8 (SEQ ID NO:2) is encoded by full length HGPRBMY8 cDNA; ACM4_CHICK (SEQ ID NO:7) represents the *Gallus gallus* (chicken)form of muscarinic acetylcholine receptor M4; YDBM_CAEEL (SEQ ID NO:8) is the *Caenorhabditis elegans* form of an orphan GPCR; 5H1A_HUMAN (SEQ ID NO:9) is the human form of the 5HT-1A receptor; 5H1A_MOUSE (SEQ ID NO:10) is the *Mus musculus* (house mouse) form of the 5HT-1A receptor; 5H1A_FUGRU (SEQ ID NO:11) represents the *Fugu rubripes* form of the 5HT-1A receptor; 5HT_LYMST (SEQ ID NO:12) is the *Lymnaea stagnalis* (great pond snail) form of the 5HT-1A receptor; A1AD_HUMAN (SEQ ID NO:13) is the human form of the alpha-1D adrenergic receptor;

A1AD_MOUSE (SEQ ID NO:14) represents the mouse form of the alpha-1D adrenergic receptor (alpha 1D-adrenoceptor); Q13675 (SEQ ID NO:15) is the human form of the alpha 1C adrenergic receptor isoform 2; Q13729 (SEQ ID NO:16) represents the human form of the alpha 1C adrenergic receptor isoform 3; O60451 is the human form of the alpha 1A adrenergic receptor isoform 4 (SEQ ID NO:17); A1AA_RAT (SEQ ID NO:18) is the Rattus norvegicus (Norway rat) form of the alpha-1A adrenergic receptor; O54913 (SEQ ID NO:19) is the Mus musculus (house mouse) form of the alpha 1A-adrenergic receptor; A1AA_BOVIN (SEQ ID NO:20) represents the Bos taurus (bovine) form of the alpha-1A adrenergic receptor; A1AA_CANFA (SEQ ID NO:21) is the Canis familiaris (dog) form of the alpha-1A adrenergic receptor; A1AA_RABIT (SEQ ID NO:22) represents the Oryctolagus cuniculus (rabbit) form of the alpha-1A adrenergic receptor; A1AA_HUMAN (SEQ ID NO:23) is the human form of the alpha-1A adrenergic receptor; A1AA_ORYLA (SEQ ID NO:24) is the Oryzias latipes (japanese medaka) form of the alpha-1A adrenergic receptor (MAR1); and O96716 (SEQ ID NO:25) represents the Branchiostoma lanceolatum (amphioxus) form of the dopamine D1/beta receptor; and O75963 (SEQ ID NO:40) is the human form of the G-protein coupled receptor RE2.

[0042] Figure 7 shows the expression profiling of the novel human orphan GPCR, HGPRBMY8, as described in Example 3.

[0043] Figure 8 shows the brain-specific expression profiling of the novel human orphan GPCR, HGPRBMY8, as described in Example 4.

[0044] Figure 9 shows the multiple sequence alignment of HGPRBMY8 and other potential SNP variants (amino acid alignment). The blackened areas represent identical amino acids and the grey highlighted areas represent similar amino acids. As shown in Figure 9, the sequences are aligned according to their amino acids, where: AL390879 (SEQ ID NO:41), AX148250 (SEQ ID NO:42), and AX080495 (SEQ ID NO:43) are compared to HGPRBMY8 (SEQ ID NO:2).

[0045] Figures 10A-D shows the multiple sequence alignment of HGPRBMY8 and other potential SNP variants (nucleic acid alignment). The blackened areas represent identical amino acids and the grey highlighted areas represent similar amino acids. As shown in Figure 10, the sequences are aligned

according to their nucleic acids, where: AX080495 (SEQ ID NO:44); AL390879 (SEQ ID NO:45), AX148250 (SEQ ID NO:46), and are compared to HGPRBMY8 (SEQ ID NO:47).

[0046] Figure 11 shows the FACS profile of an untransfected CHO-NFAT/CRE cell line.

[0047] Figure 12 shows that overexpression of HGPRBMY8 constitutively couples through the NFAT/CRE Response Element.

[0048] Figure 13 shows the FACS profile for the untransfected cAMP Response Element.

[0049] Figure 14 shows the overexpression of HGPRBMY8 results in coupling through the cAMP Response Element.

[0050] Figure 15A-D shows the localization of expressed HGPRBMY8 to the cell surface.

[0051] Figure 16A-D shows representative transfected CHO-NFAT/CRE cell lines with intermediate and high beta lactamase expression levels useful in screens to identify HGPRBMY8 agonists and/or antagonists.

[0052] Figure 17 shows the expression profiling of the novel human orphan GPCR, HGPRBMY8, as described in Example 8 and Table 1.

[0053] Figures 18A-B show the polynucleotide sequence (SEQ ID NO:48) and deduced amino acid sequence (SEQ ID NO:49) of the human G-protein coupled receptor, HGPRBMY8, comprising, or alternatively consisting of, one or more of the predicted polynucleotide polymorphic loci, in addition to, the encoded polypeptide polymorphic loci of the present invention for this particular protein.

DETAILED DESCRIPTION OF THE PRESENT INVENTION

[0054] The present invention provides a novel isolated polynucleotide and encoded polypeptide, the expression of which is high in brain. This novel polypeptide is termed herein HGPRBMY8, an acronym for "Human G-Protein coupled Receptor BMY8". HGPRBMY8 is also referred to as GPCR58 and GPCR84.

Definitions

[0055] The HGPRBMY8 polypeptide (or protein) refers to the amino acid sequence of substantially purified HGPRBMY8, which may be obtained from any

species, preferably mammalian, and more preferably, human, and from a variety of sources, including natural, synthetic, semi-synthetic, or recombinant. Functional fragments of the HGPRBMY8 polypeptide are also embraced by the present invention.

[0056] An "agonist" refers to a molecule which, when bound to the HGPRBMY8 polypeptide, or a functional fragment thereof, increases or prolongs the duration of the effect of the HGPRBMY8 polypeptide. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules that bind to and modulate the effect of HGPRBMY8 polypeptide. An antagonist refers to a molecule which, when bound to the HGPRBMY8 polypeptide, or a functional fragment thereof, decreases the amount or duration of the biological or immunological activity of HGPRBMY8 polypeptide. "Antagonists" may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules that decrease or reduce the effect of HGPRBMY8 polypeptide.

[0057] "Nucleic acid sequence", as used herein, refers to an oligonucleotide, nucleotide, or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or anti-sense strand. By way of non-limiting example, fragments include nucleic acid sequences that are greater than 20-60 nucleotides in length, and preferably include fragments that are at least 70-100 nucleotides, or which are at least 1000 nucleotides or greater in length.

[0058] Similarly, "amino acid sequence" as used herein refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragments or portions thereof, and to naturally occurring or synthetic molecules. Amino acid sequence fragments are typically from about 5 to about 30, preferably from about 5 to about 15 amino acids in length and retain the biological activity or function of the HGPRBMY8 polypeptide.

[0059] Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms, such as "polypeptide" or "protein" are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited

protein molecule. In addition, the terms HGPRBMY8 polypeptide and HGPRBMY8 protein are used interchangeably herein to refer to the encoded product of the HGPRBMY8 nucleic acid sequence of the present invention.

[0060] A "variant" of the HGPRBMY8 polypeptide refers to an amino acid sequence that is altered by one or more amino acids. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine. More rarely, a variant may have "non-conservative" changes, e.g., replacement of a glycine with a tryptophan. Minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing functional biological or immunological activity may be found using computer programs well known in the art, for example, DNASTAR software.

[0061] An "allele" or "allelic sequence" is an alternative form of the HGPRBMY8 nucleic acid sequence. Alleles may result from at least one mutation in the nucleic acid sequence and may yield altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given gene, whether natural or recombinant, may have none, one, or many allelic forms. Common mutational changes, which give rise to alleles, are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

[0062] "Altered" nucleic acid sequences encoding HGPRBMY8 polypeptide include nucleic acid sequences containing deletions, insertions and/or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent HGPRBMY8 polypeptide. Altered nucleic acid sequences may further include polymorphisms of the polynucleotide encoding the HGPRBMY8 polypeptide; such polymorphisms may or may not be readily detectable using a particular oligonucleotide probe. The encoded protein may also contain deletions, insertions, or substitutions of amino acid residues, which produce a silent change and result in a functionally equivalent HGPRBMY8 protein. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility,

hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological activity of HGPRBMY8 protein is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid; positively charged amino acids may include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

[0063] "Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide ("oligo") linked via an amide bond, similar to the peptide backbone of amino acid residues. PNAs typically comprise oligos of at least 5 nucleotides linked via amide bonds. PNAs may or may not terminate in positively charged amino acid residues to enhance binding affinities to DNA. Such amino acids include, for example, lysine and arginine, among others. These small molecules stop transcript elongation by binding to their complementary strand of nucleic acid (P.E. Nielsen et al., 1993, Anticancer Drug Des., 8:53-63). PNA may be pegylated to extend their lifespan in the cell where they preferentially bind to complementary single stranded DNA and RNA.

[0064] "Oligonucleotides" or "oligomers" refer to a nucleic acid sequence, preferably comprising contiguous nucleotides, of at least about 6 nucleotides to about 60 nucleotides, preferably at least about 8 to 10 nucleotides in length, more preferably at least about 12 nucleotides in length e.g., about 15 to 35 nucleotides, or about 15 to 25 nucleotides, or about 20 to 35 nucleotides, which can be typically used in PCR amplification assays, hybridization assays, or in microarrays. It will be understood that the term oligonucleotide is substantially equivalent to the terms primer, probe, or amplimer, as commonly defined in the art. It will also be appreciated by those skilled in the pertinent art that a longer oligonucleotide probe, or mixtures of probes, e.g., degenerate probes, can be used to detect longer, or more complex, nucleic acid sequences, for example, genomic DNA. In such cases, the probe may comprise at least 20-200 nucleotides, preferably, at least 30-100 nucleotides, more preferably, 50-100 nucleotides.

[0065] "Amplification" refers to the production of additional copies of a nucleic acid sequence and is generally carried out using polymerase chain reaction (PCR) technologies, which are well known and practiced in the art (see, D.W. Dieffenbach and G.S. Dveksler, 1995, PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview, NY).

[0066] "Microarray" is an array of distinct polynucleotides or oligonucleotides synthesized on a substrate, such as paper, nylon, or other type of membrane; filter; chip; glass slide; or any other type of suitable solid support.

[0067] The term "antisense" refers to nucleotide sequences, and compositions containing nucleic acid sequences, which are complementary to a specific DNA or RNA sequence. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. Antisense (i.e., complementary) nucleic acid molecules include PNA and may be produced by any method, including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes, which block either transcription or translation. The designation "negative" is sometimes used in reference to the antisense strand, and "positive" is sometimes used in reference to the sense strand.

[0068] The term "consensus" refers to the sequence that reflects the most common choice of base or amino acid at each position among a series of related DNA, RNA or protein sequences. Areas of particularly good agreement often represent conserved functional domains.

[0069] A "deletion" refers to a change in either nucleotide or amino acid sequence and results in the absence of one or more nucleotides or amino acid residues. By contrast, an insertion (also termed "addition") refers to a change in a nucleotide or amino acid sequence that results in the addition of one or more nucleotides or amino acid residues, as compared with the naturally occurring molecule. A substitution refers to the replacement of one or more nucleotides or amino acids by different nucleotides or amino acids.

[0070] A "derivative" nucleic acid molecule refers to the chemical modification of a nucleic acid encoding, or complementary to, the encoded

HGPRBMY8 polypeptide. Such modifications include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A nucleic acid derivative encodes a polypeptide, which retains the essential biological and/or functional characteristics of the natural molecule. A derivative polypeptide is one, which is modified by glycosylation, pegylation, or any similar process that retains the biological and/or functional or immunological activity of the polypeptide from which it is derived.

[0071] The term "biologically active", i.e., functional, refers to a protein or polypeptide or fragment thereof having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HGPRBMY8, or any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells, for example, to generate antibodies, and to bind with specific antibodies.

[0072] The term "hybridization" refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

[0073] The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary G and C bases and between complementary A and T bases. The hydrogen bonds may be further stabilized by base stacking interactions. The two complementary nucleic acid sequences hydrogen bond in an anti-parallel configuration. A hybridization complex may be formed in solution (e.g., C_ot or R_ot analysis), or between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., membranes, filters, chips, pins, or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been affixed).

[0074] The terms "stringency" or "stringent conditions" refer to the conditions for hybridization as defined by nucleic acid composition, salt and temperature. These conditions are well known in the art and may be altered to identify and/or detect identical or related polynucleotide sequences in a sample. A variety of equivalent conditions comprising either low, moderate, or high stringency depend on factors such as the length and nature of the sequence (DNA, RNA, base composition), reaction milieu (in solution or immobilized on a solid substrate), nature of the target

nucleic acid (DNA, RNA, base composition), concentration of salts and the presence or absence of other reaction components (e.g., formamide, dextran sulfate and/or polyethylene glycol) and reaction temperature (within a range of from about 5°C below the melting temperature of the probe to about 20°C to 25°C below the melting temperature). One or more factors may be varied to generate conditions, either low or high stringency that is different from but equivalent to the aforementioned conditions.

[0075] As will be understood by those of skill in the art, the stringency of hybridization may be altered in order to identify or detect identical or related polynucleotide sequences. As will be further appreciated by the skilled practitioner, the melting temperature, T_m, can be approximated by the formulas as known in the art, depending on a number of parameters, such as the length of the hybrid or probe in number of nucleotides, or hybridization buffer ingredients and conditions (see, for example, T. Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982 and J. Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; Current Protocols in Molecular Biology, Eds. F.M. Ausubel et al., Vol. 1, "Preparation and Analysis of DNA", John Wiley and Sons, Inc., 1994-1995, Suppls. 26, 29, 35 and 42; pp. 2.10.7- 2.10.16; G.M. Wahl and S. L. Berger (1987; Methods Enzymol. 152:399-407); and A.R. Kimmel, 1987; Methods of Enzymol. 152:507-511). As a general guide, T_m decreases approximately 1°C –1.5°C with every 1% decrease in sequence homology. Also, in general, the stability of a hybrid is a function of sodium ion concentration and temperature. Typically, the hybridization reaction is initially performed under conditions of low stringency, followed by washes of varying, but higher stringency. Reference to hybridization stringency, e.g., high, moderate, or low stringency, typically relates to such washing conditions.

[0076] Thus, by way of non-limiting example, "high stringency" refers to conditions that permit hybridization of those nucleic acid sequences that form stable hybrids in 0.018M NaCl at about 65°C (i.e., if a hybrid is not stable in 0.018M NaCl at about 65°C, it will not be stable under high stringency conditions). High stringency conditions can be provided, for instance, by hybridization in 50% formamide, 5x Denhardt's solution, 5xSSPE (saline sodium phosphate EDTA) (1x SSPE buffer

comprises 0.15 M NaCl, 10 mM Na_2HPO_4, 1 mM EDTA), (or 1x SSC buffer containing 150 mM NaCl, 15 mM Na_3 citrate • 2 H_2O, pH 7.0), 0.2% SDS at about 42°C, followed by washing in 1x SSPE (or saline sodium citrate, SSC) and 0.1% SDS at a temperature of at least about 42°C, preferably about 55°C, more preferably about 65°C.

[0077] "Moderate stringency" refers, by non-limiting example, to conditions that permit hybridization in 50% formamide, 5x Denhardt's solution, 5xSSPE (or SSC), 0.2% SDS at 42°C (to about 50°C), followed by washing in 0.2x SSPE (or SSC) and 0.2% SDS at a temperature of at least about 42°C, preferably about 55°C, more preferably about 65°C.

[0078] "Low stringency" refers, by non-limiting example, to conditions that permit hybridization in 10% formamide, 5x Denhardt's solution, 6xSSPE (or SSC), 0.2% SDS at 42°C, followed by washing in 1x SSPE (or SSC) and 0.2% SDS at a temperature of about 45°C, preferably about 50°C.

[0079] For additional stringency conditions, see T. Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982). It is to be understood that the low, moderate and high stringency hybridization / washing conditions may be varied using a variety of ingredients, buffers and temperatures well known to and practiced by the skilled artisan.

[0080] The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base pairing. For example, the sequence "A-G-T" binds to the complementary sequence "T-C-A". Complementarity between two single-stranded molecules may be "partial", in which only some of the nucleic acids bind, or it may be complete when total complementarity exists between single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, as well as in the design and use of PNA molecules.

[0081] The term "homology" refers to a degree of complementarity. There may be partial homology or complete homology, wherein complete homology is

equivalent to identity. A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to using the functional term "substantially homologous". The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (e.g., Southern or Northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence or probe to the target sequence under conditions of low stringency. Nonetheless, conditions of low stringency do not permit non-specific binding; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% identity). In the absence of non-specific binding, the probe will not hybridize to the second non-complementary target sequence.

[0082] Those having skill in the art will know how to determine percent identity between or among sequences using, for example, algorithms such as those based on the CLUSTALW computer program (J.D. Thompson et al., 1994, *Nucleic Acids Research*, 2(22):4673-4680), or FASTDB, (Brutlag et al., 1990, *Comp. App. Biosci.*, 6:237-245), as known in the art. Although the FASTDB algorithm typically does not consider internal non-matching deletions or additions in sequences, i.e., gaps, in its calculation, this can be corrected manually to avoid an overestimation of the % identity. CLUSTALW, however, does take sequence gaps into account in its identity calculations.

[0083] A "composition" comprising a given polynucleotide sequence refers broadly to any composition containing the given polynucleotide sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequence (SEQ ID NO:1) encoding HGPRBMY8 polypeptide (SEQ ID NO:2), or fragments thereof, may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be in association with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be

employed in an aqueous solution containing salts (e.g., NaCl), detergents or surfactants (e.g., SDS) and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, and the like).

[0084] The term "substantially purified" refers to nucleic acid sequences or amino acid sequences that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% to 85% free, and most preferably 90% or greater free from other components with which they are naturally associated.

[0085] The term "sample", or "biological sample", is meant to be interpreted in its broadest sense. A biological sample suspected of containing nucleic acid encoding HGPRBMY8 protein, or fragments thereof, or HGPRBMY8 protein itself, may comprise a body fluid, an extract from cells or tissue, chromosomes isolated from a cell (e.g., a spread of metaphase chromosomes), organelle, or membrane isolated from a cell, a cell, nucleic acid such as genomic DNA (in solution or bound to a solid support such as for Southern analysis), RNA (in solution or bound to a solid support such as for Northern analysis), cDNA (in solution or bound to a solid support), a tissue, a tissue print and the like.

[0086] "Transformation" refers to a process by which exogenous DNA enters and changes a recipient cell. It may occur under natural or artificial conditions using various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and partial bombardment. Such "transformed" cells include stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome. Transformed cells also include those cells, which transiently express the inserted DNA or RNA for limited periods of time.

[0087] The term "mimetic" refers to a molecule, the structure of which is developed from knowledge of the structure of HGPRBMY8 protein, or portions thereof, and as such, is able to effect some or all of the actions of HGPRBMY8 protein.

[0088] The term "portion" with regard to a protein (as in "a portion of a given protein") refers to fragments or segments of that protein. The fragments may range in size from four or five amino acid residues to the entire amino acid sequence minus one amino acid. Thus, a protein "comprising at least a portion of the amino acid sequence of SEQ ID NO: 2" encompasses the full-length human HGPRBMY8 polypeptide, and fragments thereof.

[0089] The term "antibody" refers to intact molecules as well as fragments thereof, such as Fab, $F(ab')_2$, Fv, or Fc, which are capable of binding an epitopic or antigenic determinant. Antibodies that bind to HGPRBMY8 polypeptides can be prepared using intact polypeptides or fragments containing small peptides of interest or prepared recombinantly for use as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal can be derived from the transition of RNA or synthesized chemically, and can be conjugated to a carrier protein, if desired. Commonly used carriers that are chemically coupled to peptides include, but are not limited to, bovine serum albumin (BSA), keyhole limpet hemocyanin (KLH), and thyroglobulin. The coupled peptide is then used to immunize the animal (e.g, a mouse, a rat, or a rabbit).

[0090] The term "humanized" antibody refers to antibody molecules in which amino acids have been replaced in the non-antigen binding regions in order to more closely resemble a human antibody, while still retaining the original binding capability, e.g., as described in U.S. Patent No. 5,585,089 to C.L. Queen et al.

[0091] The term "antigenic determinant" refers to that portion of a molecule that makes contact with a particular antibody (i.e., an epitope). When a protein or fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to a given region or three-dimensional structure on the protein; these regions or structures are referred to an antigenic determinants. An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

[0092] The terms "specific binding" or "specifically binding" refer to the interaction between a protein or peptide and a binding molecule, such as an agonist,

an antagonist, or an antibody. The interaction is dependent upon the presence of a particular structure (i.e., an antigenic determinant or epitope) of the protein that is recognized by the binding molecule. For example, if an antibody is specific for epitope "A", the presence of a protein containing epitope A (or free, unlabeled A) in a reaction containing labeled "A" and the antibody will reduce the amount of labeled A bound to the antibody.

[0093] The term "correlates with expression of a polynucleotide" indicates that the detection of the presence of ribonucleic acid that is similar to SEQ ID NO:1 by Northern analysis is indicative of the presence of mRNA encoding HGPRBMY8 polypeptide (SEQ ID NO:2) in a sample and thereby correlates with expression of the transcript from the polynucleotide encoding the protein.

[0094] An alteration in the polynucleotide of SEQ ID NO:1 comprises any alteration in the sequence of the polynucleotides encoding HGPRBMY8 polypeptide, including deletions, insertions, and point mutations that may be detected using hybridization assays. Included within this definition is the detection of alterations to the genomic DNA sequence which encodes HGPRBMY8 polypeptide (e.g., by alterations in the pattern of restriction fragment length polymorphisms capable of hybridizing to SEQ ID NO:1), the inability of a selected fragment of SEQ ID NO:1 to hybridize to a sample of genomic DNA (e.g., using allele-specific oligonucleotide probes), and improper or unexpected hybridization, such as hybridization to a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HGPRBMY8 polypeptide (e.g., using fluorescent *in situ* hybridization (FISH) to metaphase chromosome spreads).

DESCRIPTION OF THE PRESENT INVENTION

[0095] The present invention provides a novel human member of the G-protein coupled receptor (GPCR) family (HGPRBMY8). Based on sequence homology, the protein HGPRBMY8 is a novel human GPCR. This protein sequence has been predicted to contain seven transmembrane domains which is a characteristic structural feature of GPCRs. HGPRBMY8 belongs to the "class A" of GPCR superfamily and is closely related to adrenergic and serotonin receptors based on

sequence similarity. Class A is the largest sub-family of the GPCR superfamily. This particular orphan GPCR is expressed highly in brain.

[0096] HGPRBMY8 polypeptides and polynucleotides are useful for diagnosing diseases related to over- or under- expression of HGPRBMY8 proteins by identifying mutations in the HGPRBMY8 gene using HGPRBMY8 probes, or by determining HGPRBMY8 protein or mRNA expression levels. HGPRBMY8 polypeptides are also useful for screening compounds, which affect activity or function of the protein. The invention encompasses the polynucleotide encoding the HGPRBMY8 polypeptide and the use of the HGPRBMY8 polynucleotide or polypeptide, or composition thereof, in the screening, diagnosis, treatment, or prevention of disorders associated with aberrant or uncontrolled cellular growth and/or function, such as neoplastic diseases (e.g., cancers and tumors), with particular regard to diseases or disorders related to the brain, e.g. neurological disorders.

[0097] Nucleic acids encoding human HGPRBMY8 according to the present invention were first identified from the human genomic data available from GenBank (Accession No: AC016468).

[0098] In one of its embodiments, the present invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:2 as shown in Figure 1. The HGPRBMY8 polypeptide is 508 amino acids in length and shares amino acid sequence homology with the GPCR RE2. The HGPRBMY8 polypeptide (SEQ ID NO:2) shares 24.3 % identity and 33.6 % similarity with over 400 amino acids of the GPCR RE2 sequence, wherein "similar" amino acids are those which have the same/ similar physical properties and in many cases, the function is conserved with similar residues. For example, amino acids Lysine and Arginine are similar; while residues such as Proline and Cysteine, which do not share any physical properties, are considered dissimilar. The HGPRBMY8 polypeptide shares 28.01% identity and 38.33% similarity with the *Fugu rubripes* 5-Hydroxytryptamine 1a-Alpha Receptor (5H1A_FUGRU; Acc. No.:O42385); 25.3% identity and 37.23% similarity with the human 5-Hydroxytryptamine 1a-Alpha Receptor (5H1A_HUMAN; Acc. No.:P08908); 27.56% idenity and 37.56% similarity with the Mus musculus 5-Hydroxytryptamine 1a-Alpha Receptor (5H1A_MOUSE; Acc. No.:Q64264,

Q60956); 25.46% identity and 37.05% similarity with the *Lymnaea stagnalis* 5-hydroxytryptamine receptor (5HT_LYMST; Acc. No.:Q25414); 23.67% identity and 33.19% similarity with the Bos taurus Alpha-1A adrenergic receptor (A1AA_BOVIN; Acc. No.: P18130); 26.21% identity and 36.9% similarity with the *Canis familiaris* Alpha-1A adrenergic receptor (A1AA_CANFA; Acc. No.: O77621); 29.47% identity and 41.05% similarity with the human Alpha-1A adrenergic receptor (A1AA_HUMAN; Acc. No.: P35348); 31.65% identity and 42.29% similarity with the *Oryzias latipes* Alpha-1A adrenergic receptor (A1AA_ORYLA; Acc. No.:Q91175); 30% identity and 41.32% similarity with the *Oryctolagus cuniculus* Alpha-1A adrenergic receptor (A1AA_RABIT; Acc. No.: O02824); 24.82% identity and 34.43% similarity with the *Rattus norvegicus* Alpha-1A adrenergic receptor (A1AA_RAT; Acc. No.:P43140); 29.79% identity and 41.19% similarity with the human Alpha-1D adrenergic receptor (A1AD_HUMAN; Acc. No.: P25100); 29.2% identity and 40.57% similarity with the Mus musculus Alpha-1D adrenergic receptor (A1AD_MOUSE; Acc. No.:P97714, Q61619); 23.33% identity and 31.97% similarity with the *Gallus gallus* muscarinic acetylcholine receptor M4 (ACM4_CHICK; Acc. No.:P17200); 30.53% identity and 41.58% similarity with the *Mus musculus* Alpha-1A adrenergic receptor (O54913; Acc. No.:O54913); 29.47% identity and 41.05% similarity with the human Alpha-1A adrenergic receptor isoform 4 (O60451; Acc. No.:O60451); 23.59% identity and 32.82% similarity with the human G-protein coupled receptor RE2 (O75963; Acc. No.:O75963); 23.99% identity and 31.81% similarity with the Branchiostoma lanceolatum dopamine D1/Beta receptor (O96716; Acc. No.:O96716); 29.21% identity and 40.79% similarity with the human Alpha 1C adrenergic receptor isoform 2 (Q13675; Acc. No.:Q13675); 24.87% identity and 34.52% similarity with the human Alpha 1C adrenergic receptor isoform 3 (Q13729; Acc. No.:Q13729); and 21.49% identity and 32.023% similarity with the *Caenorhabditis elegans* probable G protein coupled receptor F01E11.5 (YDBM_CAEEL; Acc. No.:Q19084).

[0099] Variants of the HGPRBMY8 polypeptide are also encompassed by the present invention. A preferred HGPRBMY8 variant has at least 75 to 80%, more preferably at least 85 to 90%, and even more preferably at least 90% amino acid

sequence identity to the amino acid sequence claimed herein, and which retains at least one biological, immunological, or other functional characteristic or activity of the HGPRBMY8 polypeptide. Most preferred is a variant having at least 95% amino acid sequence identity to that of SEQ ID NO:2. For example, Figures 9 and 10 show multiple sequence alignments of HGPRBMY8 and single nucleotide polymorphism (SNP) variants. Highlighted are the differences in sequence.

[0100] In a preferred embodiment, polynucleotide and polypeptide polymorphisms are shown in Figure 18A-B. The standard one-letter abbreviation for amino acids is used to illustrate the deduced amino acid sequence. The polynucleotide sequence contains a sequence of 1527 nucleotides (SEQ ID NO:48), encoding a polypeptide of 508 amino acids (SEQ ID NO:49). The polynucleotide polymorphic sites are represented by an "N", in bold. The polypeptide polymorphic sites are represented by an "X", and underlined. The present invention encompasses the polynucleotide at nucleotide position 370 as being either a "T" or a "G", the polynucleotide at nucleotide position 1055 as being either a "A" or a "G", the polynucleotide at nucleotide position 1192 as being either a "G" or a "A", the polynucleotide at nucleotide position 1193 as being either a "C" or a "A", and the polynucleotide at nucleotide position 1194 as being either a "T" or a "G" of Figures 18A-B (SEQ ID NO:48), in addition to any combination thereof. The present invention also encompasses the polypeptide at amino acid position 124 as being either a "Leu" or a "Val", the polypeptide at amino acid position 352 as being either a "Asp" or a "Gly", and the polypeptide at amino acid position 398 as being either a "Ala" or an "Lys" of Figures 18A-B (SEQ ID NO:49).

[0101] These polymorphisms are useful as genetic markers for any study that attempts to look for linkage between HGPRBMY8 and a disease or disease state related to this polypeptide.

[0102] In preferred embodiments, the following single nucleotide polymorphism polynucleotides are encompassed by the present invention:
CACCATTGTCTTGGTGTCAGT (SEQ ID NO:50),
CACCATTGTCGTGGTGTCAGT (SEQ ID NO:51),
GGTGAAGATGACATGGAGTTT (SEQ ID NO:52),

GGTGAAGATGGCATGGAGTTT (SEQ ID NO:53),
GTGCAAAGCTGCTAAAGTGAT (SEQ ID NO:54),
GTGCAAAGCTACTAAAGTGAT (SEQ ID NO:55),
TGCAAAGCTGCTAAAGTGATC (SEQ ID NO:56),
TGCAAAGCTGATAAAGTGATC (SEQ ID NO:57)
GCAAAGCTGCTAAAGTGATCT (SEQ ID NO:58), and/or
GCAAAGCTGCGAAAGTGATCT (SEQ ID NO:59). Polypeptides encoded by these polynucleotides are also provided.

[0103] The predicted 'T' to 'G' polynucleotide polymorphism located at nucleic acid 370 of SEQ ID NO:1 is a missense mutation resulting in a change in an encoding amino acid from 'L' to 'V' at amino acid position 124 of SEQ ID NO:2.

[0104] The predicted 'A' to 'G' polynucleotide polymorphism located at nucleic acid 1055 of SEQ ID NO:1 is a missense mutation resulting in a change in an encoding amino acid from 'D' to 'G' at amino acid position 352 of SEQ ID NO:2.

[0105] The predicted 'G' to 'A' polynucleotide polymorphism located at nucleic acid 1192 of SEQ ID NO:1 is a missense mutation resulting in a change in an encoding amino acid from 'A' to 'T' at amino acid position 398 of SEQ ID NO:2.

[0106] The predicted 'C' to 'A' polynucleotide polymorphism located at nucleic acid 1193 of SEQ ID NO:1 is a missense mutation resulting in a change in an encoding amino acid from 'A' to 'D' at amino acid position 398 of SEQ ID NO:2.

[0107] The predicted 'T' to 'G' polynucleotide polymorphism located at nucleic acid 1194 of SEQ ID NO:1 is a silent mutation and does not result in a change in amino acid.

[0108] However, taken together the predicted 'G' to 'A' polynucleotide polymorphism located at nucleic acid 1192, the predicted 'C' to 'A' polynucleotide polymorphism located at nucleic acid 1193, and the predicted 'T' to 'G' polynucleotide polymorphism located at nucleic acid 1194 of SEQ ID NO:1 represent a missense mutations resulting in a change in an encoding amino acid from 'A' to 'K' at amino acid position 398 of SEQ ID NO:2.

[0109] The present invention relates to isolated nucleic acid molecules comprising, or alternatively, consisting of all or a portion of the variant allele of the

human HGPRBMY8 G-protein coupled receptor gene (e.g., wherein reference or wildtype human HGPRBMY8 G-protein coupled receptor gene is exemplified by SEQ ID NO:1). Preferred portions are at least 10, preferably at least 20, preferably at least 40, preferably at least 100, contiguous polynucleotides comprising anyone of the human HGPRBMY8 G-protein coupled receptor gene alleles described herein and exemplified in Figures 10A-D.

[0110] In one embodiment, the invention relates to a method for predicting the likelihood that an individual will have a disorder associated with the reference allele at nucleotide position 370, 1055, 1192, 1193, and/or 1194 of SEQ ID NO:1 (or diagnosing or aiding in the diagnosis of such a disorder) comprising the steps of obtaining a DNA sample from an individual to be assessed and determining the nucleotide present at position 370, 1055, 1192, 1193, and/or 1194 of SEQ ID NO:1. The presence of the variant allele at this position indicates that the individual has a greater likelihood of having a disorder associated therewith than an individual having the reference allele at that position, or a greater likelihood of having more severe symptoms.

[0111] Conversely, the invention relates to a method for predicting the likelihood that an individual will have a disorder associated with the variant allele at nucleotide position 370, 1055, 1192, 1193, and/or 1194 of SEQ ID NO:1 (or diagnosing or aiding in the diagnosis of such a disorder) comprising the steps of obtaining a DNA sample from an individual to be assessed and determining the nucleotide present at position 370, 1055, 1192, 1193, and/or 1194 of SEQ ID NO:1. The presence of the variant allele at this position indicates that the individual has a greater likelihood of having a disorder associated therewith than an individual having the reference allele at that position, or a greater likelihood of having more severe symptoms.

[0112] The present invention further relates to isolated proteins or polypeptides comprising, or alternatively, consisting of all or a portion of the encoded variant amino acid sequence of the human HGPRBMY8 G-protein coupled receptor polypeptide (e.g., wherein reference or wildtype human HGPRBMY8 G-protein coupled receptor polypeptide is exemplified by SEQ ID NO:2). Preferred portions are

at least 10, preferably at least 20, preferably at least 40, preferably at least 100, contiguous polypeptides and comprises any one of the amino acid variant alleles of the human HGPRBMY8 G-protein coupled receptor polypeptide exemplified in Figures 18A-B, or a portion of SEQ ID NO:49. Alternatively, preferred portions are at least 10, preferably at least 20, preferably at least 40, preferably at least 100, contiguous polypeptides and comprises any one of the amino acid reference alleles of the human HGPRBMY8 G-protein coupled receptor protein exemplified in Figures 18A-B, or a portion of SEQ ID NO:49. The invention further relates to isolated nucleic acid molecules encoding such polypeptides or proteins, as well as to antibodies that bind to such proteins or polypeptides.

[0113] In another embodiment, the present invention encompasses polynucleotides, which encode the HGPRBMY8 polypeptide. Accordingly, any nucleic acid sequence, which encodes the amino acid sequence of HGPRBMY8 polypeptide, can be used to produce recombinant molecules that express HGPRBMY8 protein. In a particular embodiment, the present invention encompasses the HGPRBMY8 polynucleotide comprising the nucleic acid sequence of SEQ ID NO:1 as shown in Figure 1. More particularly, the present invention provides the HGPRBMY8 clone. More particularly, the present invention provides the HGPRBMY8 clone, deposited at the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209 on January 24, 2001 and under ATCC Accession No. PTA-2966 according to the terms of the Budapest Treaty.

[0114] As will be appreciated by the skilled practitioner in the art, the degeneracy of the genetic code results in the production of a number of nucleotide sequences encoding HGPRBMY8 polypeptide. Some of the sequences bear minimal homology to the nucleotide sequences of any known and naturally occurring gene. Accordingly, the present invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring HGPRBMY8, and all such variations are to be considered as being specifically disclosed.

[0115] Although nucleotide sequences which encode HGPRBMY8 polypeptide and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring HGPRBMY8 polypeptide under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HGPRBMY8 polypeptide, or its derivatives, which possess a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide/polypeptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HGPRBMY8 polypeptide, and its derivatives, without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

[0116] The present invention also encompasses production of DNA sequences, or portions thereof, which encode the HGPRBMY8 polypeptide, and its derivatives, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents that are well known and practiced by those in the art. Moreover, synthetic chemistry and other known techniques may be used to introduce mutations into a sequence encoding HGPRBMY8 polypeptide, or any fragment thereof.

[0117] In preferred embodiments, the present invention encompasses a polynucleotide lacking the initiating start codon, in addition to, the resulting encoded polypeptide of HGPRBMY8. Specifically, the present invention encompasses the polynucleotide corresponding to nucleotides 4 thru 1524 of SEQ ID NO:1, and the polypeptide corresponding to amino acids 2 thru 508 of SEQ ID NO:2. Also encompassed are recombinant vectors comprising said encoding sequence, and host cells comprising said vector.

[0118] Also encompassed by the present invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequence of HGPRBMY8, such as that shown in SEQ ID NO:1, under various conditions of stringency. Hybridization conditions are typically based on the melting temperature

(T_m) of the nucleic acid binding complex or probe (see, G.M. Wahl and S.L. Berger, 1987; Methods Enzymol., 152:399-407 and A.R. Kimmel, 1987; Methods of Enzymol., 152:507-511), and may be used at a defined stringency. For example, included in the present invention are sequences capable of hybridizing under moderately stringent conditions to the HGPRBMY8 sequence of SEQ ID NO:1 and other sequences which are degenerate to those which encode HGPRBMY8 polypeptide (e.g., as a non-limiting example: prewashing solution of 2X SSC, 0.5% SDS, 1.0mM EDTA, pH 8.0, and hybridization conditions of 50°C, 5XSSC, overnight.

[0119] The nucleic acid sequence encoding the HGPRBMY8 protein may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one method, which may be employed, is restriction-site PCR, which utilizes universal primers to retrieve unknown sequence adjacent to a known locus (G. Sarkar, 1993, PCR Methods Applic., 2:318-322). In particular, genomic DNA is first amplified in the presence of primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

[0120] Inverse PCR may also be used to amplify or extend sequences using divergent primers based on a known region or sequence (T. Triglia et al., 1988, Nucleic Acids Res., 16:8186). The primers may be designed using OLIGO 4.06 Primer Analysis software (National Biosciences Inc.; Plymouth, MN), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72°C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

[0121] Another method which may be used is capture PCR which involves PCR amplification of DNA fragments adjacent to a known sequence in human and

yeast artificial chromosome (YAC) DNA (M. Lagerstrom et al., 1991, PCR Methods Applic., 1:111-119). In this method, multiple restriction enzyme digestions and ligations may also be used to place an engineered double-stranded sequence into an unknown portion of the DNA molecule before performing PCR. J.D. Parker et al. (1991; *Nucleic Acids Res.*, 19:3055-3060) provide another method which may be used to retrieve unknown sequences. In addition, PCR, nested primers, and PROMOTERFINDER libraries can be used to walk genomic DNA (Clontech, Palo Alto, CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

[0122] When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Also, random-primed libraries are preferable, since they will contain more sequences, which contain the 5' regions of genes. The use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into the 5' and 3' non-transcribed regulatory regions.

[0123] The embodiments of the present invention can be practiced using methods for DNA sequencing which are well known and generally available in the art. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical Corp.; Cleveland, OH), Taq polymerase (PE Biosystems; Gaithersburg, MD), thermostable T7 polymerase (Amersham Pharmacia Biotechnology; Piscataway, NJ), or combinations of recombinant polymerases and proofreading exonucleases such as the ELONGASE Amplification System marketed by Life Technologies (Rockville, MD). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton; Reno, NV), Peltier Thermal Cycler (PTC200; MJ Research; Watertown, MA) and the ABI Catalyst and 373 and 377 DNA sequencers (PE Biosystems; Gaithersburg, MD).

[0124] Commercially available capillary electrophoresis systems may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for

electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, PE Biosystems; Gaithersburg, MD) and the entire process -- from loading of samples to computer analysis and electronic data display -- may be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA, which might be present in limited amounts in a particular sample.

[0125] In another embodiment of the present invention, polynucleotide sequences or fragments thereof which encode HGPRBMY8 polypeptide, or peptides thereof, may be used in recombinant DNA molecules to direct the expression of HGPRBMY8 polypeptide product, or fragments or functional equivalents thereof, in appropriate host cells. Because of the inherent degeneracy of the genetic code, other DNA sequences, which encode substantially the same or a functionally equivalent amino acid sequence, may be produced and these sequences may be used to clone and express HGPRBMY8 protein.

[0126] As will be appreciated by those having skill in the art, it may be advantageous to produce HGPRBMY8 polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

[0127] The nucleotide sequence of the present invention can be engineered using methods generally known in the art in order to alter HGPRBMY8 polypeptide-encoding sequences for a variety of reasons, including, but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce

mutations, and the like.

[0128] In another embodiment of the present invention, natural, modified, or recombinant nucleic acid sequences encoding HGPRBMY8 polypeptide may be ligated to a heterologous sequence to encode a fusion protein. For example, for screening peptide libraries for inhibitors of HGPRBMY8 activity, it may be useful to encode a chimeric HGPRBMY8 protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the HGPRBMY8 protein-encoding sequence and the heterologous protein sequence, so that HGPRBMY8 protein may be cleaved and purified away from the heterologous moiety.

[0129] In another embodiment, sequences encoding HGPRBMY8 polypeptide may be synthesized in whole, or in part, using chemical methods well known in the art (see, for example, M.H. Caruthers et al., 1980, Nucl. Acids Res. Symp. Ser., 215-223 and T. Horn et al., 1980, Nucl. Acids Res. Symp. Ser., 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of HGPRBMY8 polypeptide, or a fragment or portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (J.Y. Roberge et al., 1995, Science, 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (PE Biosystems; Gaithersburg, MD).

[0130] The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (e.g., T. Creighton, 1983, Proteins, Structures and Molecular Principles, W.H. Freeman and Co., New York, NY), by reversed-phase high performance liquid chromatography, or other purification methods as are known in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure; Creighton, *supra*). In addition, the amino acid sequence of HGPRBMY8 polypeptide or any portion thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

[0131] To express a biologically active HGPRBMY8 polypeptide or peptide,

the nucleotide sequences encoding HGPRBMY8 polypeptide, or functional equivalents, may be inserted into an appropriate expression vector, i.e., a vector, which contains the necessary elements for the transcription and translation of the inserted coding sequence.

[0132] Methods, which are well known to those skilled in the art, may be used to construct expression vectors containing sequences encoding HGPRBMY8 polypeptide and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described in J. Sambrook et al., 1989, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y. and in F.M. Ausubel et al., 1989, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY.

[0133] A variety of expression vector/ host systems may be utilized to contain and express sequences encoding HGPRBMY8 polypeptide. Such expression vector/host systems include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus (CaMV) and tobacco mosaic virus (TMV)), or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The host cell employed is not limiting to the present invention.

[0134] "Control elements" or "regulatory sequences" are those non-translated regions of the vector, e.g., enhancers, promoters, 5' and 3' untranslated regions, which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene; La Jolla, CA) or PSPORT1 plasmid (Life Technologies; Rockville, MD), and the like, may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers

derived from the genomes of plant cells (e.g., heat shock, RUBISCO; and storage protein genes), or from plant viruses (e.g., viral promoters or leader sequences), may be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding HGPRBMY8, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

[0135] In bacterial systems, a number of expression vectors may be selected, depending upon the use intended for the expressed HGPRBMY8 product. For example, when large quantities of expressed protein are needed for the induction of antibodies, vectors, which direct high level expression of fusion proteins that are readily purified, may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene; La Jolla, CA), in which the sequence encoding HGPRBMY8 polypeptide may be ligated into the vector in-frame with sequences for the amino-terminal Met and the subsequent 7 residues of ß-galactosidase, so that a hybrid protein is produced; pIN vectors (see, G. Van Heeke and S.M. Schuster, 1989, J. Biol. Chem., 264:5503-5509); and the like. pGEX vectors (Promega, Madison, WI) may also be used to express foreign polypeptides, as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can be easily purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

[0136] In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. (For reviews, see F.M. Ausubel et al., *supra*, and Grant et al., 1987, *Methods Enzymol.*, 153:516-544).

[0137] Should plant expression vectors be desired and used, the expression of sequences encoding HGPRBMY8 polypeptide may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from

TMV (N. Takamatsu, 1987, EMBO J., 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO, or heat shock promoters, may be used (G. Coruzzi et al., 1984, EMBO J., 3:1671-1680; R. Broglie et al., 1984, Science, 224:838-843; and J. Winter et al., 1991, Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, S. Hobbs or L.E. Murry, In: McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

[0138] An insect system may also be used to express HGPRBMY8 polypeptide. For example, in one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The sequences encoding HGPRBMY8 polypeptide may be cloned into a non-essential region of the virus such as the polyhedrin gene and placed under control of the polyhedrin promoter. Successful insertion of HGPRBMY8 polypeptide will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which the HGPRBMY8 polypeptide product may be expressed (E.K. Engelhard et al., 1994, Proc. Nat. Acad. Sci., 91:3224-3227).

[0139] In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HGPRBMY8 polypeptide may be ligated into an adenovirus transcription/ translation complex containing the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing HGPRBMY8 polypeptide in infected host cells (J. Logan and T. Shenk, 1984, Proc. Natl. Acad. Sci., 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

[0140] Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HGPRBMY8 polypeptide. Such signals include

the ATG initiation codon and adjacent sequences. In cases where sequences encoding HGPRBMY8 polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals, including the ATG initiation codon, should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system that is used, such as those described in the literature (D. Scharf et al., 1994, Results Probl. Cell Differ., 20:125-162).

[0141] Moreover, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells having specific cellular machinery and characteristic mechanisms for such post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and W138) are available from the American Type Culture Collection (ATCC), American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, and may be chosen to ensure the correct modification and processing of the foreign protein.

[0142] For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express HGPRBMY8 protein may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same, or on a separate, vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched cell culture medium before they are switched to selective medium. The purpose of the selectable marker is to confer

resistance to selection, and its presence allows the growth and recovery of cells, which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

[0143] Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the Herpes Simplex Virus thymidine kinase (HSV TK), (M. Wigler et al., 1977, Cell, 11:223-32) and adenine phosphoribosyltransferase (I. Lowy et al., 1980, Cell, 22:817-23) genes which can be employed in tk^- or $aprt^-$ cells, respectively. Also, anti-metabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr, which confers resistance to methotrexate (M. Wigler et al., 1980, Proc. Natl. Acad. Sci., 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and G-418 (F. Colbere-Garapin et al., 1981, J. Mol. Biol., 150:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (S.C. Hartman and R.C. Mulligan, 1988, Proc. Natl. Acad. Sci., 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as the anthocyanins, ß-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, which are widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression that is attributable to a specific vector system (C.A. Rhodes et al., 1995, *Methods Mol. Biol.*, 55:121-131).

[0144] Although the presence or absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the desired gene of interest may need to be confirmed. For example, if the nucleic acid sequence encoding HGPRBMY8 polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences encoding HGPRBMY8 polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HGPRBMY8 polypeptide under the control of a single promoter. Expression of the marker gene in response to induction

or selection usually indicates co-expression of the tandem gene.

[0145] Alternatively, host cells, which contain the nucleic acid, sequence encoding HGPRBMY8 polypeptide and which express HGPRBMY8 polypeptide product may be identified by a variety of procedures known to those having skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques, including membrane, solution, or chip based technologies, for the detection and/or quantification of nucleic acid or protein.

[0146] The presence of polynucleotide sequences encoding HGPRBMY8 polypeptide can be detected by DNA-DNA or DNA-RNA hybridization, or by amplification using probes or portions or fragments of polynucleotides encoding HGPRBMY8 polypeptide. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers, based on the sequences encoding HGPRBMY8 polypeptide, to detect transformants containing DNA or RNA encoding HGPRBMY8 polypeptide.

[0147] A wide variety of labels and conjugation techniques are known and employed by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HGPRBMY8 polypeptide include oligo-labeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HGPRBMY8 polypeptide, or any portions or fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase, such as T7, T3, or SP(6) and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits (e.g., Amersham Pharmacia Biotech, Promega and U.S. Biochemical Corp.). Suitable reporter molecules or labels which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

[0148] Furthermore, in yet another embodiment, G-protein coupled receptor-

encoding polynucleotide sequences can be used to purify a molecule or compound in a sample, wherein the molecule or compound specifically binds to the polynucleotide, comprising: a) combining the G-protein coupled receptor-encoding polynucleotide, or fragment thereof, under conditions to allow specific binding; b) detecting specific binding between the G-protein coupled receptor-encoding polynucleotide and the molecule or compound; c) recovering the bound polynucleotide; and d) separating the polynucleotide from the molecule or compound, thereby obtaining a purified molecule or compound.

[0149] Host cells transformed with nucleotide sequences encoding HGPRBMY8 protein, or fragments thereof, may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/ or the vector used. As will be understood by those having skill in the art, expression vectors containing polynucleotides which encode HGPRBMY8 protein may be designed to contain signal sequences which direct secretion of the HGPRBMY8 protein through a prokaryotic or eukaryotic cell membrane. Other constructions may be used to join nucleic acid sequences encoding HGPRBMY8 protein to nucleotide sequence encoding a polypeptide domain, which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals; protein A domains that allow purification on immobilized immunoglobulin; and the domain utilized in the FLAGS extension/ affinity purification system (Immunex Corp., Seattle, WA). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, CA) between the purification domain and HGPRBMY8 protein may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing HGPRBMY8 and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMAC (immobilized metal ion affinity chromatography) as described by J. Porath et al., 1992, Prot. Exp. Purif., 3:263-281, while the enterokinase cleavage site provides a means for purifying from the fusion

protein. For a discussion of suitable vectors for fusion protein production, see D.J. Kroll et al., 1993; DNA Cell Biol., 12:441-453.

[0150] In addition to recombinant production, fragments of HGPRBMY8 polypeptide may be produced by direct peptide synthesis using solid-phase techniques (J. Merrifield, 1963, J. Am. Chem. Soc., 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using ABI 431A Peptide Synthesizer (PE Biosystems; Gaithersburg, MD). Various fragments of HGPRBMY8 polypeptide can be chemically synthesized separately and then combined using chemical methods to produce the full-length molecule.

[0151] Human artificial chromosomes (HACs) may be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid vector. HACs are linear microchromosomes which may contain DNA sequences of 10K to 10M in size, and contain all of the elements that are required for stable mitotic chromosome segregation and maintenance (see, J.J. Harrington et al., 1997, *Nature Genet.*, 15:345-355). HACs of 6 to 10M are constructed and delivered via conventional delivery methods (e.g., liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes.

Diagnostic Assays

[0152] A variety of protocols for detecting and measuring the expression of HGPRBMY8 polypeptide using either polyclonal or monoclonal antibodies specific for the protein are known and practiced in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive with two non-interfering epitopes on the HGPRBMY8 polypeptide is preferred, but a competitive binding assay may also be employed. These and other assays are described in the art as represented by the publication of R. Hampton et al., 1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, MN and D.E. Maddox et al., 1983; J. Exp. Med., 158:1211-1216).

[0153] This invention also relates to the use of HGPRBMY8 polynucleotides as diagnostic reagents. Detection of a mutated form of the HGPRBMY8 gene

associated with a dysfunction provides a diagnostic tool that can add to or define a diagnosis of a disease or susceptibility to a disease which results from under-expression, over-expression, or altered expression of HGPRBMY8. Individuals carrying mutations in the HGPRBMY8 gene may be detected at the DNA level by a variety of techniques.

[0154] Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Hybridizing amplified DNA to labeled HGPRBMY8 polynucleotide sequences can identify point mutations. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. See, e.g., Myers et al., Science (1985) 230:1242. Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method. See Cotton et al., Proc. Natl. Acad. Sci., USA (1985) 85:43297-4401. In another embodiment, an array of oligonucleotides probes comprising HGPRBMY8 nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M. Chee et al., Science, 274:610-613, 1996).

[0155] The diagnostic assays offer a process for diagnosing or determining, for example, a susceptibility to infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2 through detection of a mutation in the HGPRBMY8 gene by the methods described. The invention also provides diagnostic assays for determining or monitoring susceptibility to the

following conditions, diseases, or disorders: HIV infections; asthma; allergies; obesity; anorexia; bulimia; ulcers; acute heart failure; hypotension; hypertension; angina pectoris; myocardial infarction; urinary retention; osteoporosis; benign prostatic hypertrophy; cancers; brain-related disorders; Parkinson's disease; neuropathic pain; immune; metabolic; cardiovascular; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome; Sydenham chorea; major depressive disorder; and obsessive-compulsive disorder (OCD). Movement type diseases, disorders, or conditions may be targeted in particular since HGPRBMY8 is expressed in the caudate nucleus of the brain.

[0156] In addition, infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2, as well as, conditions, diseases, or disorders such as, HIV infections; asthma; allergies; obesity; anorexia; bulimia; ulcers; acute heart failure; hypotension; hypertension; angina pectoris; myocardial infarction; urinary retention; osteoporosis; benign prostatic hypertrophy; cancers; brain-related disorders; Parkinson's disease; neuropathic pain; immune; metabolic; cardiovascular; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, can be diagnosed by methods comprising determining from a sample derived from a subject having an abnormally decreased or increased level of HGPRBMY8 polypeptide or HGPRBMY8 mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantification of polynucleotides, such as, for example, PCR, RT-PCR, RNAse protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as an HGPRBMY8, in a sample derived from a host are well known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

[0157] In another of its aspects, the present invention relates to a diagnostic

kit for a disease or susceptibility to a disease, particularly infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2, as well as, conditions, diseases, or disorders such as, HIV infections; asthma; allergies; obesity; anorexia; bulimia; ulcers; acute heart failure; hypotension; hypertension; angina pectoris; myocardial infarction; urinary retention; osteoporosis; benign prostatic hypertrophy; cancers; brain-related disorders; Parkinson's disease; neuropathic pain; immune; metabolic; cardiovascular; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, which comprises:

(a) an HGPRBMY8 polynucleotide, preferably the nucleotide sequence of SEQ ID NO: 1, or a fragment thereof; or

(b) a nucleotide sequence complementary to that of (a); or

(c) an HGPRBMY8 polypeptide, preferably the polypeptide of SEQ ID NO: 2, or a fragment thereof; or

(d) an antibody to an HGPRBMY8 polypeptide, preferably to the polypeptide of SEQ ID NO: 2, or combinations thereof.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component and instructions are frequently included.

[0158] The GPCR polynucleotides which may be used in the diagnostic assays according to the present invention include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify HGPRBMY8-encoding nucleic acid expression in biopsied tissues in which expression (or under- or overexpression) of the HGPRBMY8 polynucleotide may be correlated with disease. The diagnostic assays may be used to distinguish between the absence, presence, and excess expression of HGPRBMY8, and to monitor regulation of HGPRBMY8 polynucleotide levels during therapeutic treatment or intervention.

[0159] In a related aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HGPRBMY8 polypeptide, or closely related molecules, may be used to identify nucleic acid sequences which encode HGPRBMY8 polypeptide. The specificity of

the probe, whether it is made from a highly specific region, e.g., about 8 to 10 contiguous nucleotides in the 5' regulatory region, or a less specific region, e.g., especially in the 3' coding region, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding HGPRBMY8 polypeptide, alleles thereof, or related sequences.

[0160] Probes may also be used for the detection of related sequences, and should preferably contain at least 50% of the nucleotides, most optimally 15-35 nucleotides, encoding the HGPRBMY8 polypeptide. The hybridization probes of this invention may be DNA or RNA and may be derived from the nucleotide sequence of SEQ ID NO:1, or from genomic sequence including promoter, enhancer elements, and introns of the naturally occurring HGPRBMY8 protein.

[0161] Methods for producing specific hybridization probes for DNA encoding the HGPRBMY8 polypeptide include the cloning of a nucleic acid sequence that encodes the HGPRBMY8 polypeptide, or HGPRBMY8 derivatives, into vectors for the production of mRNA probes. Such vectors are known in the art, commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of detector/ reporter groups, e.g., radionuclides such as ^{32}P or ^{35}S, or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/ biotin coupling systems, and the like.

[0162] The polynucleotide sequence encoding the HGPRBMY8 polypeptide, or fragments thereof, may be used for the diagnosis of disorders associated with expression of HGPRBMY8. Examples of such disorders or conditions are described for "Therapeutics". The polynucleotide sequence encoding the HGPRBMY8 polypeptide may be used in Southern or Northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or chip assays utilizing fluids or tissues from patient biopsies to detect the status of, e.g., levels or overexpression of HGPRBMY8, or to detect altered HGPRBMY8 expression. Such qualitative or quantitative methods are well known in the art.

[0163] In a particular aspect, the nucleotide sequence encoding the

HGPRBMY8 polypeptide may be useful in assays that detect activation or induction of various neoplasms or cancers, particularly those mentioned *supra*. The nucleotide sequence encoding the HGPRBMY8 polypeptide may be labeled by standard methods, and added to a fluid or tissue sample from a patient, under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the biopsied or extracted sample is significantly altered from that of a comparable control sample, the nucleotide sequence has hybridized with nucleotide sequence present in the sample, and the presence of altered levels of nucleotide sequence encoding the HGPRBMY8 polypeptide in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or in monitoring the treatment of an individual patient.

[0164] To provide a basis for the diagnosis of disease associated with expression of HGPRBMY8, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, which encodes the HGPRBMY8 polypeptide, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject (patient) values is used to establish the presence of disease.

[0165] Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that which is observed in a normal individual. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

[0166] With respect to cancer, the presence of an abnormal amount of

transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier, thereby preventing the development or further progression of the cancer.

[0167] Additional diagnostic uses for oligonucleotides designed from the nucleic acid sequence encoding the HGPRBMY8 polypeptide may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically, or produced from a recombinant source. Oligomers will preferably comprise two nucleotide sequences, one with sense orientation (5'→3') and another with antisense (3'→5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantification of closely related DNA or RNA sequences.

[0168] Methods suitable for quantifying the expression of HGPRBMY8 include radiolabeling or biotinylating nucleotides, co-amplification of a control nucleic acid, and standard curves onto which the experimental results are interpolated (P.C. Melby et al., 1993, J. Immunol. Methods, 159:235-244; and C. Duplaa et al., 1993, Anal. Biochem., 229-236). The speed of quantifying multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantification.

Therapeutic Assays

[0169] The HGPRBMY8 polypeptide (SEQ ID NO:2) shares homology with somatostatin-type receptors. The HGPRBMY8 protein may play a role in neurological disorders, and/or in cell cycle regulation, and/or in cell signaling. The HGPRBMY8 protein may further be involved in neoplastic, cardiovascular, and immunological disorders.

[0170] In one embodiment of the present invention, the HGPRBMY8 protein

may play a role in neoplastic disorders. An antagonist or inhibitor of the HGPRBMY8 polypeptide may be administered to an individual to prevent or treat a neoplastic disorder. Such disorders may include, but are not limited to, adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, and teratocarcinoma, and particularly, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. In a related aspect, an antibody which specifically binds to HGPRBMY8 may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express the HGPRBMY8 polypeptide.

[0171] In another embodiment of the present invention, an antagonist or inhibitory agent of the HGPRBMY8 polypeptide may be administered to an individual to prevent or treat an immunological disorder. Such disorders may include, but are not limited to, AIDS, Addison's disease, adult respiratory distress syndrome, allergies, anemia, asthma, atherosclerosis, bronchitis, cholecystitis, Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythema nodosum, atrophic gastritis, glomerulonephritis, gout, Graves' disease, hypereosinophilia, irritable bowel syndrome, lupus erythematosus, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome, and autoimmune thyroiditis; complications of cancer, hemodialysis, extracorporeal circulation; viral, bacterial, fungal, parasitic, protozoal, and helminthic infections and trauma.

[0172] In a preferred embodiment of the present invention, an antagonist or inhibitory agent of the HGPRBMY8 polypeptide may be administered to an individual to prevent or treat a neurological disorder, particularly since HGPRBMY8 is highly expressed in the brain. Such disorders may include, but are not limited to, akathesia, Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, bipolar disorder, catatonia, cerebral neoplasms, dementia, depression, Down's syndrome, tardive dyskinesia, dystonias, epilepsy, Huntington's disease, multiple sclerosis,

Parkinson's disease, paranoid psychoses, schizophrenia, and Tourette's disorder.

[0173] In preferred embodiments, the HGPRBMY8 polynucleotides and polypeptides, including agonists, antagonists, and fragments thereof, are useful for modulating intracellular cAMP associated signaling pathways.

[0174] In another embodiment of the present invention, an expression vector containing the complement of the polynucleotide encoding HGPRBMY8 polypeptide may be administered to an individual to treat or prevent a neoplastic disorder, including, but not limited to, the types of cancers and tumors described above.

[0175] In yet another embodiment of the present invention, an expression vector containing the complement of the polynucleotide encoding HGPRBMY8 polypeptide may be administered to an individual to treat or prevent an immune disorder, including, but not limited to, the types of immune disorders described above.

[0176] In a preferred embodiment of the present invention, an expression vector containing the complement of the polynucleotide encoding HGPRBMY8 polypeptide may be administered to an individual to treat or prevent a neurological disorder, including, but not limited to, the types of disorders described above.

[0177] In another embodiment, the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the present invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

[0178] Antagonists or inhibitors of the HGPRBMY8 polypeptide of the present invention may be produced using methods which are generally known in the art. For example, the HGPRBMY8 transfected CHO-NFAT/CRE cell lines of the present invention are useful for the identification of agonists and antagonists of the HGPRBMY8 polypeptide. Representative uses of these cell lines would be their inclusion in a method of identifying HGPRBMY8 agonists and antagonists.

Preferably, the cell lines are useful in a method for identifying a compound that modulates the biological activity of the HGPRBMY8 polypeptide, comprising the steps of (a) combining a candidate modulator compound with a host cell expressing the HGPRBMY8 polypeptide having the sequence as set forth in SEQ ID NO:2; and (b) measuring an effect of the candidate modulator compound on the activity of the expressed HGPRBMY8 polypeptide. Representative vectors expressing the HGPRBMY8 polypeptide are referenced herein (e.g., pcDNA3.1 hygro ™) or otherwise known in the art.

[0179] The cell lines are also useful in a method of screening for a compounds that is capable of modulating the biological activity of HGPRBMY8 polypeptide, comprising the steps of: (a) determining the biological activity of the HGPRBMY8 polypeptide in the absence of a modulator compound; (b) contacting a host cell expression the HGPRBMY8 polypeptide with the modulator compound; and (c) determining the biological activity of the HGPRBMY8 polypeptide in the presence of the modulator compound; wherein a difference between the activity of the HGPRBMY8 polypeptide in the presence of the modulator compound and in the absence of the modulator compound indicates a modulating effect of the compound. Additional uses for these cell lines are described herein or otherwise known in the art. In particular, purified HGPRBMY8 protein, or fragments thereof, can be used to produce antibodies, or to screen libraries of pharmaceutical agents, to identify those which specifically bind HGPRBMY8.

[0180] Antibodies specific for HGPRBMY8 polypeptide, or immunogenic peptide fragments thereof, can be generated using methods that have long been known and conventionally practiced in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments, and fragments produced by an Fab expression library. Neutralizing antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

[0181] The present invention also encompasses the polypeptide sequences that intervene between each of the predicted HGPRBMY8 transmembrane domains. Since these regions are solvent accessible either extracellularly or intracellularly, they are particularly useful for designing antibodies specific to each region. Such

antibodies may be useful as antagonists or agonists of the HGPRBMY8 full-length polypeptide and may modulate its activity.

[0182] The following serve as non-limiting examples of peptides or fragments that may be used to generate antibodies:

MTSTCTNSTRESNSSHTCMPLSKMPISLAHGIIRST (SEQ ID NO:26)

QRKPQLLQVTNRF (SEQ ID NO:27)

WPLNS (SEQ ID NO:28)

DRYLSIIHPLSYPSKMTQRR (SEQ ID NO:29)

GQAAFDERNALCSMIWGASPSYT (SEQ ID NO:30)

CAARRQHALLYNVKRHSLEVRVKDCVENEDEEGAEKKEEFQDESEFRRQ HEGEVKAKEGRMEAKDGSLKAKEGSTGTSESSVEAGSEEVRESSTVA SDGSMEGKEGSTKVEENSMKADKGRTEVNQCSIDLGEDDMEFGEDDI NFSEDDVEAVNIPESLPPSRRNSNSNPPLPRCYQCKAAK (SEQ ID NO:31)

AVLAVWVDVETQVPQ (SEQ ID NO:32)

YGYMHKTIKKEIQDMLKKFFCKEKPPKEDSHPDLPGTEGGTEGKIVPSYD SATFP (SEQ ID NO:33)

[0183] The present invention also encompasses the polypeptide sequences that intervene between each of the predicted HGPRBMY8 transmembrane domains. Since these regions are solvent accessible either extracellularly or intracellularly, they are particularly useful for designing antibodies specific to each region. Such antibodies may be useful as antagonists or agonists of the HGPRBMY8 full-length polypeptide and may modulate its activity.

[0184] In preferred embodiments, the following N-terminal HGPRBMY8 TM1-2 intertransmembrane domain deletion polypeptides are encompassed by the present invention: Q1-F13, R2-F13, K3-F13, P4-F13, Q5-F13, L6-F13, and/or L7-F13 of SEQ ID NO:27. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these N-terminal HGPRBMY8 TM1-2 intertransmembrane domain deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0185] In preferred embodiments, the following C-terminal HGPRBMY8

TM1-2 intertransmembrane domain deletion polypeptides are encompassed by the present invention: Q1-F13, Q1-R12, Q1-N11, Q1-T10, Q1-V9, Q1-Q8, and/or Q1-L7 of SEQ ID NO:27. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these C-terminal HGPRBMY8 TM1-2 intertransmembrane domain deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0186] In preferred embodiments, the following N-terminal HGPRBMY8 TM3-4 intertransmembrane domain deletion polypeptides are encompassed by the present invention: D1-R20, R2-R20, Y3-R20, L4-R20, S5-R20, I6-R20, I7-R20, H8-R20, P9-R20, L10-R20, S11-R20, Y12-R20, P13-R20, and/or S14-R20 of SEQ ID NO:29. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these N-terminal HGPRBMY8 TM3-4 intertransmembrane domain deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0187] In preferred embodiments, the following C-terminal HGPRBMY8 TM3-4 intertransmembrane domain deletion polypeptides are encompassed by the present invention: D1-R20, D1-R19, D1-Q18, D1-T17, D1-M16, D1-K15, D1-S14, D1-P13, D1-Y12, D1-S11, D1-L10, D1-P9, D1-H8, and/or D1-I7 of SEQ ID NO:29. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these C-terminal HGPRBMY8 TM3-4 intertransmembrane domain deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0188] In preferred embodiments, the following N-terminal HGPRBMY8 TM4-5 intertransmembrane domain deletion polypeptides are encompassed by the present invention: G1-T23, Q2-T23, A3-T23, A4-T23, F5-T23, D6-T23, E7-T23, R8-T23, N9-T23, A10-T23, L11-T23, C12-T23, S13-T23, M14-T23, I15-T23, W16-T23, and/or G17-T23 of SEQ ID NO:30. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these N-terminal HGPRBMY8 TM4-5 intertransmembrane domain deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0189] In preferred embodiments, the following C-terminal HGPRBMY8 TM4-5 intertransmembrane domain deletion polypeptides are encompassed by the present invention: G1-T23, G1-Y22, G1-S21, G1-P20, G1-S19, G1-A18, G1-G17, G1-W16, G1-I15, G1-M14, G1-S13, G1-C12, G1-L11, G1-A10, G1-N9, G1-R8, and/or G1-E7 of SEQ ID NO:30. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these C-terminal HGPRBMY8 TM4-5 intertransmembrane domain deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0190] In preferred embodiments, the following N-terminal HGPRBMY8 TM5-6 intertransmembrane domain deletion polypeptides are encompassed by the present invention: C1-K182, A2-K182, A3-K182, R4-K182, R5-K182, Q6-K182, H7-K182, A8-K182, L9-K182, L10-K182, Y11-K182, N12-K182, V13-K182, K14-K182, R15-K182, H16-K182, S17-K182, L18-K182, E19-K182, V20-K182, R21-K182, V22-K182, K23-K182, D24-K182, C25-K182, V26-K182, E27-K182, N28-K182, E29-K182, D30-K182, E31-K182, E32-K182, G33-K182, A34-K182, E35-K182, K36-K182, K37-K182, E38-K182, E39-K182, F40-K182, Q41-K182, D42-K182, E43-K182, S44-K182, E45-K182, F46-K182, R47-K182, R48-K182, Q49-K182, H50-K182, E51-K182, G52-K182, E53-K182, V54-K182, K55-K182, A56-K182, K57-K182, E58-K182, G59-K182, R60-K182, M61-K182, E62-K182, A63-K182, K64-K182, D65-K182, G66-K182, S67-K182, L68-K182, K69-K182, A70-K182, K71-K182, E72-K182, G73-K182, S74-K182, T75-K182, G76-K182, T77-K182, S78-K182, E79-K182, S80-K182, S81-K182, V82-K182, E83-K182, A84-K182, G85-K182, S86-K182, E87-K182, E88-K182, V89-K182, R90-K182, E91-K182, S92-K182, S93-K182, T94-K182, V95-K182, A96-K182, S97-K182, D98-K182, G99-K182, S100-K182, M101-K182, E102-K182, G103-K182, K104-K182, E105-K182, G106-K182, S107-K182, T108-K182, K109-K182, V110-K182, E111-K182, E112-K182, N113-K182, S114-K182, M115-K182, K116-K182, A117-K182, D118-K182, K119-K182, G120-K182, R121-K182, T122-K182, E123-K182, V124-K182, N125-K182, Q126-K182, C127-K182, S128-K182, I129-K182, D130-K182, L131-K182, G132-K182, E133-K182, D134-K182, D135-K182, M136-K182, E137-

K182, F138-K182, G139-K182, E140-K182, D141-K182, D142-K182, I143-K182, N144-K182, F145-K182, S146-K182, E147-K182, D148-K182, D149-K182, V150-K182, E151-K182, A152-K182, V153-K182, N154-K182, I155-K182, P156-K182, E157-K182, S158-K182, L159-K182, P160-K182, P161-K182, S162-K182, R163-K182, R164-K182, N165-K182, S166-K182, N167-K182, S168-K182, N169-K182, P170-K182, P171-K182, L172-K182, P173-K182, R174-K182, C175-K182, and/or Y176-K182 of SEQ ID NO:31. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these N-terminal HGPRBMY8 TM5-6 intertransmembrane domain deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0191] In preferred embodiments, the following C-terminal HGPRBMY8 TM5-6 intertransmembrane domain deletion polypeptides are encompassed by the present invention: C1-K182, C1-A181, C1-A180, C1-K179, C1-C178, C1-Q177, C1-Y176, C1-C175, C1-R174, C1-P173, C1-L172, C1-P171, C1-P170, C1-N169, C1-S168, C1-N167, C1-S166, C1-N165, C1-R164, C1-R163, C1-S162, C1-P161, C1-P160, C1-L159, C1-S158, C1-E157, C1-P156, C1-I155, C1-N154, C1-V153, C1-A152, C1-E151, C1-V150, C1-D149, C1-D148, C1-E147, C1-S146, C1-F145, C1-N144, C1-I143, C1-D142, C1-D141, C1-E140, C1-G139, C1-F138, C1-E137, C1-M136, C1-D135, C1-D134, C1-E133, C1-G132, C1-L131, C1-D130, C1-I129, C1-S128, C1-C127, C1-Q126, C1-N125, C1-V124, C1-E123, C1-T122, C1-R121, C1-G120, C1-K119, C1-D118, C1-A117, C1-K116, C1-M115, C1-S114, C1-N113, C1-E112, C1-E111, C1-V110, C1-K109, C1-T108, C1-S107, C1-G106, C1-E105, C1-K104, C1-G103, C1-E102, C1-M101, C1-S100, C1-G99, C1-D98, C1-S97, C1-A96, C1-V95, C1-T94, C1-S93, C1-S92, C1-E91, C1-R90, C1-V89, C1-E88, C1-E87, C1-S86, C1-G85, C1-A84, C1-E83, C1-V82, C1-S81, C1-S80, C1-E79, C1-S78, C1-T77, C1-G76, C1-T75, C1-S74, C1-G73, C1-E72, C1-K71, C1-A70, C1-K69, C1-L68, C1-S67, C1-G66, C1-D65, C1-K64, C1-A63, C1-E62, C1-M61, C1-R60, C1-G59, C1-E58, C1-K57, C1-A56, C1-K55, C1-V54, C1-E53, C1-G52, C1-E51, C1-H50, C1-Q49, C1-R48, C1-R47, C1-F46, C1-E45, C1-S44, C1-E43, C1-D42, C1-Q41, C1-F40, C1-E39, C1-E38, C1-K37, C1-K36, C1-E35, C1-A34, C1-G33, C1-E32, C1-

E31, C1-D30, C1-E29, C1-N28, C1-E27, C1-V26, C1-C25, C1-D24, C1-K23, C1-V22, C1-R21, C1-V20, C1-E19, C1-L18, C1-S17, C1-H16, C1-R15, C1-K14, C1-V13, C1-N12, C1-Y11, C1-L10, C1-L9, C1-A8, and/or C1-H7 of SEQ ID NO:31. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these C-terminal HGPRBMY8 TM5-6 intertransmembrane domain deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0192] In preferred embodiments, the following N-terminal HGPRBMY8 TM6-7 intertransmembrane domain deletion polypeptides are encompassed by the present invention: A1-Q15, V2-Q15, L3-Q15, A4-Q15, V5-Q15, W6-Q15, V7-Q15, D8-Q15, and/or V9-Q15 of SEQ ID NO:32. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these N-terminal HGPRBMY8 TM6-7 intertransmembrane domain deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0193] In preferred embodiments, the following C-terminal HGPRBMY8 TM6-7 intertransmembrane domain deletion polypeptides are encompassed by the present invention: A1-Q15, A1-P14, A1-V13, A1-Q12, A1-T11, A1-E10, A1-V9, A1-D8, and/or A1-V7 of SEQ ID NO:32. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these C-terminal HGPRBMY8 TM6-7 intertransmembrane domain deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0194] The HGPRBMY8 polypeptide was predicted to comprise eight PKC phosphorylation sites using the Motif algorithm (Genetics Computer Group, Inc.). In vivo, protein kinase C exhibits a preference for the phosphorylation of serine or threonine residues. The PKC phosphorylation sites have the following consensus pattern: [ST]-x-[RK], where S or T represents the site of phosphorylation and 'x' an intervening amino acid residue. Additional information regarding PKC phosphorylation sites can be found in Woodget J.R., Gould K.L., Hunter T., Eur. J. Biochem. 161:177-184(1986), and Kishimoto A., Nishiyama K., Nakanishi H.,

Uratsuji Y., Nomura H., Takeyama Y., Nishizuka Y., J. Biol. Chem. 260:12492-12499(1985); which are hereby incorporated by reference herein.

[0195] In preferred embodiments, the following PKC phosphorylation site polypeptides are encompassed by the present invention: STCTNSTRESNSS (SEQ ID NO:76), QLLQVTNRFIFNL (SEQ ID NO:77), YPSKMTQRRGYLL (SEQ ID NO:78), EAKDGSLKAKEGS (SEQ ID NO:79), EGKEGSTKVEENS (SEQ ID NO:80), KVEENSMKADKGR (SEQ ID NO:81), ESLPPSRRNSNSN (SEQ ID NO:82), and/or GYMHKTIKKEIQD (SEQ ID NO:83). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of the HGPRBMY8 PKC phosphorylation site polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0196] The HGPRBMY8 polypeptide was predicted to comprise five casein kinase II phosphorylation sites using the Motif algorithm (Genetics Computer Group, Inc.). Casein kinase II (CK-2) is a protein serine/threonine kinase whose activity is independent of cyclic nucleotides and calcium. CK-2 phosphorylates many different proteins. The substrate specificity [1] of this enzyme can be summarized as follows: (1) Under comparable conditions Ser is favored over Thr.; (2) An acidic residue (either Asp or Glu) must be present three residues from the C-terminal of the phosphate acceptor site; (3) Additional acidic residues in positions +1, +2, +4, and +5 increase the phosphorylation rate. Most physiological substrates have at least one acidic residue in these positions; (4) Asp is preferred to Glu as the provider of acidic determinants; and (5) A basic residue at the N-terminal of the acceptor site decreases the phosphorylation rate, while an acidic one will increase it.

[0197] A consensus pattern for casein kinase II phosphorylations site is as follows: [ST]-x(2)-[DE], wherein 'x' represents any amino acid, and S or T is the phosphorylation site.

[0198] Additional information specific to aminoacyl-transfer RNA synthetases class-II domains may be found in reference to the following publication: Pinna L.A., Biochim. Biophys. Acta 1054:267-284(1990); which is hereby incorporated herein in its entirety.

[0199] In preferred embodiments, the following casein kinase II

phosphorylation site polypeptide is encompassed by the present invention: STCTNSTRESNSSH (SEQ ID NO:84), TGTSESSVEARGSE (SEQ ID NO:85), GKEGSTKVEENSMK (SEQ ID NO:86), DDINFSEDDVEAVN (SEQ ID NO:87), and/or PPKEDSHPDLPGTE (SEQ ID NO:88). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of this casein kinase II phosphorylation site polypeptide as an immunogenic and/or antigenic epitope as described elsewhere herein.

[0200] The HGPRBMY8 polypeptide was predicted to comprise two cAMP- and cGMP-dependent protein kinase phosphorylation site using the Motif algorithm (Genetics Computer Group, Inc.). There has been a number of studies relative to the specificity of cAMP- and cGMP-dependent protein kinases. Both types of kinases appear to share a preference for the phosphorylation of serine or threonine residues found close to at least two consecutive N-terminal basic residues.

[0201] A consensus pattern for cAMP- and cGMP-dependent protein kinase phosphorylation sites is as follows: [RK](2)-x-[ST], wherein "x" represents any amino acid, and S or T is the phosphorylation site.

[0202] Additional information specific to cAMP- and cGMP-dependent protein kinase phosphorylation sites may be found in reference to the following publication: Fremisco J.R., Glass D.B., Krebs E.G, J. Biol. Chem. 255:4240-4245(1980); Glass D.B., Smith S.B., J. Biol. Chem. 258:14797-14803(1983); and Glass D.B., El-Maghrabi M.R., Pilkis S.J., J. Biol. Chem. 261:2987-2993(1986); which is hereby incorporated herein in its entirety.

[0203] In preferred embodiments, the following cAMP- and cGMP-dependent protein kinase phosphorylation site polypeptide is encompassed by the present invention: LLYNVKRHSLEVRV (SEQ ID NO:89), and/or SLPPSRRNSNSNPP (SEQ ID NO:90). Polynucleotides encoding this polypeptide are also provided. The present invention also encompasses the use of this cAMP- and cGMP-dependent protein kinase phosphorylation site polypeptide as an immunogenic and/or antigenic epitope as described elsewhere herein.

[0204] The HGPRBMY8 polypeptide has been shown to comprise three glycosylation sites according to the Motif algorithm (Genetics Computer Group, Inc.).

As discussed more specifically herein, protein glycosylation is thought to serve a variety of functions including: augmentation of protein folding, inhibition of protein aggregation, regulation of intracellular trafficking to organelles, increasing resistance to proteolysis, modulation of protein antigenicity, and mediation of intercellular adhesion.

[0205] Asparagine glycosylation sites have the following consensus pattern, N-{P}-[ST]-{P}, wherein N represents the glycosylation site. However, it is well known that that potential N-glycosylation sites are specific to the consensus sequence Asn-Xaa-Ser/Thr. However, the presence of the consensus tripeptide is not sufficient to conclude that an asparagine residue is glycosylated, due to the fact that the folding of the protein plays an important role in the regulation of N-glycosylation. It has been shown that the presence of proline between Asn and Ser/Thr will inhibit N-glycosylation; this has been confirmed by a recent statistical analysis of glycosylation sites, which also shows that about 50% of the sites that have a proline C-terminal to Ser/Thr are not glycosylated. Additional information relating to asparagine glycosylation may be found in reference to the following publications, which are hereby incorporated by reference herein: Marshall R.D., Annu. Rev. Biochem. 41:673-702(1972); Pless D.D., Lennarz W.J., Proc. Natl. Acad. Sci. U.S.A. 74:134-138(1977); Bause E., Biochem. J. 209:331-336(1983); Gavel Y., von Heijne G., Protein Eng. 3:433-442(1990); and Miletich J.P., Broze G.J. Jr., J. Biol. Chem. 265:11397-11404(1990).

[0206] In preferred embodiments, the following asparagine glycosylation site polypeptides are encompassed by the present invention: TSTCTNSTRESNSS (SEQ ID NO:91), STRESNSSHTCMPL (SEQ ID NO:92), and/or GEDDINFSEDDVEA (SEQ ID NO:93). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of these HGPRBMY8 asparagine glycosylation site polypeptide as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0207] The HGPRBMY8 polypeptide was predicted to comprise eight N-myristoylation sites using the Motif algorithm (Genetics Computer Group, Inc.). An appreciable number of eukaryotic proteins are acylated by the covalent addition of

myristate (a C14-saturated fatty acid) to their N-terminal residue via an amide linkage. The sequence specificity of the enzyme responsible for this modification, myristoyl CoA:protein N-myristoyl transferase (NMT), has been derived from the sequence of known N-myristoylated proteins and from studies using synthetic peptides. The specificity seems to be the following: i.) The N-terminal residue must be glycine; ii.) In position 2, uncharged residues are allowed; iii.) Charged residues, proline and large hydrophobic residues are not allowed; iv.) In positions 3 and 4, most, if not all, residues are allowed; v.) In position 5, small uncharged residues are allowed (Ala, Ser, Thr, Cys, Asn and Gly). Serine is favored; and vi.) In position 6, proline is not allowed.

[0208] A consensus pattern for N-myristoylation is as follows: G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}, wherein 'x' represents any amino acid, and G is the N-myristoylation site.

[0209] Additional information specific to N-myristoylation sites may be found in reference to the following publication: Towler D.A., Gordon J.I., Adams S.P., Glaser L., Annu. Rev. Biochem. 57:69-99(1988); and Grand R.J.A., Biochem. J. 258:625-638(1989); which is hereby incorporated herein in its entirety.

[0210] In preferred embodiments, the following N-myristoylation site polypeptides are encompassed by the present invention: ISLAHGIIRSTVLVIF (SEQ ID NO:94), CSMIWGASPSYTILSV (SEQ ID NO:95), MEAKDGSLKAKEGSTG (SEQ ID NO:96), LKAKEGSTGTSESSVE (SEQ ID NO:97), KEGSTGTSESSVEARG (SEQ ID NO:98), TVASDGSMEGKEGSTK (SEQ ID NO:99), HPDLPGTEGGTEGKIV (SEQ ID NO:100), and/or LPGTEGGTEGKIVPSY (SEQ ID NO:101). The present invention also encompasses the use of these N-myristoylation site polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0211] Moreover, in confirmation of HGPRBMY8 representing a novel GPCR, the HGPRBMY8 polypeptide was predicted to comprise a G-protein coupled receptor motif using the Motif algorithm (Genetics Computer Group, Inc.). G-protein coupled receptors (also called R7G) are an extensive group of hormones, neurotransmitters, odorants and light receptors which transduce extracellular signals

by interaction with guanine nucleotide-binding (G) proteins. Some examples of receptors that belong to this family are provided as follows: 5-hydroxytryptamine (serotonin) 1A to 1F, 2A to 2C, 4, 5A, 5B, 6 and 7, Acetylcholine, muscarinic-type, M1 to M5, Adenosine A1, A2A, A2B and A3, Adrenergic alpha-1A to -1C; alpha-2A to -2D; beta-1 to –3, Angiotensin II types I and II, Bombesin subtypes 3 and 4, Bradykinin B1 and B2, c3a and C5a anaphylatoxin, Cannabinoid CB1 and CB2, Chemokines C-C CC-CKR-1 to CC-CKR-8, Chemokines C-X-C CXC-CKR-1 to CXC-CKR-4, Cholecystokinin-A and cholecystokinin-B/gastrin, Dopamine D1 to D5, Endothelin ET-a and ET-b, fMet-Leu-Phe (fMLP) (N-formyl peptide), Follicle stimulating hormone (FSH-R), Galanin, Gastrin-releasing peptide (GRP-R), Gonadotropin-releasing hormone (GNRH-R), Histamine H1 and H2 (gastric receptor I), Lutropin-choriogonadotropic hormone (LSH-R), Melanocortin MC1R to MC5R, Melatonin, Neuromedin B (NMB-R), Neuromedin K (NK-3R), Neuropeptide Y types 1 to 6, Neurotensin (NT-R), Octopamine (tyramine) from insects, Odorants, Opioids delta-, kappa- and mu-types, Oxytocin (OT-R), Platelet activating factor (PAF-R), Prostacyclin, Prostaglandin D2, Prostaglandin E2, EP1 to EP4 subtypes, Prostaglandin F2, Purinoreceptors (ATP), Somatostatin types 1 to 5, Substance-K (NK-2R), Substance-P (NK-1R), Thrombin, Thromboxane A2, Thyrotropin (TSH-R), Thyrotropin releasing factor (TRH-R), Vasopressin V1a, V1b and V2, Visual pigments (opsins and rhodopsin), Proto-oncogene mas, Caenorhabditis elegans putative receptors C06G4.5, C38C10.1, C43C3.2,T27D1.3 and ZC84.4, Three putative receptors encoded in the genome of cytomegalovirus: US27, US28, and UL33., ECRF3, a putative receptor encoded in the genome of herpesvirus saimiri.

[0212] The structure of all GPCRs are thought to be identical. They have seven hydrophobic regions, each of which most probably spans the membrane. The N-terminus is located on the extracellular side of the membrane and is often glycosylated, while the C-terminus is cytoplasmic and generally phosphorylated. Three extracellular loops alternate with three intracellular loops to link the seven transmembrane regions. Most, but not all of these receptors, lack a signal peptide. The most conserved parts of these proteins are the transmembrane regions and the first two cytoplasmic loops. A conserved acidic-Arg-aromatic triplet is present in the N-

terminal extremity of the second cytoplasmic loop and could be implicated in the interaction with G proteins.

[0213] The putative consensus sequence for GPCRs comprises the conserved triplet and also spans the major part of the third transmembrane helix, and is as follows: [GSTALIVMFYWC]-[GSTANCPDE]-{EDPKRH}-x(2)-[LIVMNQGA]-x(2)-[LIVMFT]-[GSTANC]-[LIVMFYWSTAC]-[DENH]-R-[FYWCSH]-x(2)-[LIVM], where "X" represents any amino acid.

[0214] Additional information relating to G-protein coupled receptors may be found in reference to the following publications: Strosberg A.D., Eur. J. Biochem. 196:1-10(1991); Kerlavage A.R., Curr. Opin. Struct. Biol. 1:394-401(1991); Probst W.C., Snyder L.A., Schuster D.I., Brosius J., Sealfon S.C., DNA Cell Biol. 11:1-20(1992); Savarese T.M., Fraser C.M., Biochem. J. 283:1-9(1992); Branchek T., Curr. Biol. 3:315-317(1993); Stiles G.L., J. Biol. Chem. 267:6451-6454(1992); Friell T., Kobilka B.K., Lefkowitz R.J., Caron M.G., Trends Neurosci. 11:321-324(1988); Stevens C.F., Curr. Biol. 1:20-22(1991); Sakurai T., Yanagisawa M., Masaki T., Trends Pharmacol. Sci. 13:103-107(1992); Salesse R., Remy J.J., Levin J.M., Jallal B., Garnier J., Biochimie 73:109-120(1991); Lancet D., Ben-Arie N., Curr. Biol. 3:668-674(1993); Uhl G.R., Childers S., Pasternak G., Trends Neurosci. 17:89-93(1994); Barnard E.A., Burnstock G., Webb T.E., Trends Pharmacol. Sci. 15:67-70(1994); Applebury M.L., Hargrave P.A., Vision Res. 26:1881-1895(1986); Attwood T.K., Eliopoulos E.E., Findlay J.B.C., Gene 98:153-159(1991); http://www.gcrdb.uthscsa.edu/; and http://swift.embl-heidelberg.de/7tm/.

[0215] In preferred embodiments, the following G-protein coupled receptors signature polypeptide is encompassed by the present invention: SVVSFIVIPLIVMIACYSVVF (SEQ ID NO:102). Polynucleotides encoding this polypeptide are also provided. The present invention also encompasses the use of the HGPRBMY8 G-protein coupled receptors signature polypeptide as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0216] For the production of antibodies, various hosts including goats, rabbits, sheep, rats, mice, humans, and others, can be immunized by injection with HGPRBMY8 polypeptide, or any fragment or oligopeptide thereof, which has

immunogenic properties. Depending on the host species, various adjuvants may be used to increase the immunological response. Non-limiting examples of suitable adjuvants include Freund's (complete and incomplete), mineral gels such as aluminum hydroxide or silica, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Adjuvants typically used in humans include BCG (bacilli Calmette Guérin) and *Corynebacterium parvumn*.

[0217] Preferably, the peptides, fragments, or oligopeptides used to induce antibodies to HGPRBMY8 polypeptide (i.e., immunogens) have an amino acid sequence having at least five amino acids, and more preferably, at least 7-10 amino acids. It is also preferable that the immunogens are identical to a portion of the amino acid sequence of the natural protein; they may also contain the entire amino acid sequence of a small, naturally occurring molecule. The peptides, fragments or oligopeptides may comprise a single epitope or antigenic determinant or multiple epitopes. Short stretches of HGPRBMY8 amino acids may be fused with those of another protein, such as KLH, and antibodies are produced against the chimeric molecule.

[0218] Monoclonal antibodies to HGPRBMY8 polypeptide, or immunogenic fragments thereof, may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (G. Kohler et al., 1975, Nature, 256:495-497; D. Kozbor et al., 1985, J. Immunol. Methods, 81:31-42; R.J. Cote et al., 1983, Proc. Natl. Acad. Sci. USA, 80:2026-2030; and S.P. Cole et al., 1984, Mol. Cell Biol., 62:109-120). The production of monoclonal antibodies is well known and routinely used in the art.

[0219] In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (S.L. Morrison et al., 1984, Proc. Natl. Acad. Sci. USA, 81:6851-6855; M.S. Neuberger et al., 1984, Nature, 312:604-608; and S. Takeda et al., 1985, Nature,

314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HGPRBMY8 polypeptide-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (D.R. Burton, 1991, Proc. Natl. Acad. Sci. USA, 88:11120-3). Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (R. Orlandi et al., 1989, Proc. Natl. Acad. Sci. USA, 86:3833-3837 and G. Winter et al., 1991, Nature, 349:293-299).

[0220] Antibody fragments, which contain specific binding sites for HGPRBMY8 polypeptide, may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments which can be produced by pepsin digestion of the antibody molecule and Fab fragments which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (W.D. Huse et al., 1989, Science, 254.1275-1281).

[0221] Various immunoassays can be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve measuring the formation of complexes between HGPRBMY8 polypeptide and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive with two non-interfering HGPRBMY8 polypeptide epitopes is preferred, but a competitive binding assay may also be employed (Maddox, *supra*).

[0222] Another aspect of the invention relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with HGPRBMY8 polypeptide, or a fragment thereof, adequate to produce antibody and/or T cell immune response to protect said animal from infections such as bacterial,

fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2. Yet another aspect of the invention relates to a method of inducing immunological response in a mammal which comprises, delivering HGPRBMY8 polypeptide via a vector directing expression of HGPRBMY8 polynucleotide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

[0223] A further aspect of the invention relates to an immunological/ vaccine formulation (composition) which, when introduced into a mammalian host, induces an immunological response in that mammal to an HGPRBMY8 polypeptide wherein the composition comprises an HGPRBMY8 polypeptide or HGPRBMY8 gene. The vaccine formulation may further comprise a suitable carrier. Since the HGPRBMY8 polypeptide may be broken down in the stomach, it is preferably administered parenterally (including subcutaneous, intramuscular, intravenous, intradermal, etc., injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials, and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in-water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

[0224] In an embodiment of the present invention, the polynucleotide encoding the HGPRBMY8 polypeptide, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, antisense, to the polynucleotide encoding the HGPRBMY8 polypeptide, may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HGPRBMY8 polypeptide. Thus, complementary molecules may be used to modulate

HGPRBMY8 polynucleotide and polypeptide activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or oligonucleotides, or larger fragments, can be designed from various locations along the coding or control regions of polynucleotide sequences encoding HGPRBMY8 polypeptide.

[0225] Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods, which are well known to those skilled in the art, can be used to construct recombinant vectors which will express a nucleic acid sequence that is complementary to the nucleic acid sequence encoding the HGPRBMY8 polypeptide. These techniques are described both in J. Sambrook et al., *supra* and in F.M. Ausubel et al., *supra*.

[0226] Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy". Thus for example, cells from a subject may be engineered with a polynucleotide, such as DNA or RNA, to encode a polypeptide *ex vivo*, and for example, by the use of a retroviral plasmid vector. The cells can then be introduced into the subject.

[0227] The genes encoding the HGPRBMY8 polypeptide can be turned off by transforming a cell or tissue with an expression vector that expresses high levels of an HGPRBMY8 polypeptide-encoding polynucleotide, or a fragment thereof. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and even longer if appropriate replication elements are designed to be part of the vector system.

[0228] Modifications of gene expression can be obtained by designing antisense molecules or complementary nucleic acid sequences (DNA, RNA, or PNA), to the control, 5', or regulatory regions of the gene encoding the HGPRBMY8 polypeptide, (e.g., signal sequence, promoters, enhancers, and introns). Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved

using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described (see, for example, J.E. Gee et al., 1994, In: B.E. Huber and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY). The antisense molecule or complementary sequence may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

[0229] Ribozymes, i.e., enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Suitable examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HGPRBMY8 polypeptide.

[0230] Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

[0231] Complementary ribonucleic acid molecules and ribozymes according to the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. Such methods include techniques for chemically synthesizing oligonucleotides, for example, solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding HGPRBMY8. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP. Alternatively, the cDNA constructs that constitutively or inducibly synthesize complementary RNA can be introduced into cell lines, cells, or tissues.

[0232] RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/ or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl, rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytosine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

[0233] Many methods for introducing vectors into cells or tissues are available and are equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods, which are well known in the art.

[0234] Any of the therapeutic methods described above may be applied to any individual in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

[0235] A further embodiment of the present invention embraces the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, diluent, or excipient, for any of the above-described therapeutic uses and effects. Such pharmaceutical compositions may comprise HGPRBMY8 nucleic acid, polypeptide, or peptides, antibodies to HGPRBMY8 polypeptide, mimetics, agonists, antagonists, or inhibitors of HGPRBMY8 polypeptide or polynucleotide. The compositions may be administered alone, or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, hormones, or biological response modifiers.

[0236] The pharmaceutical compositions for use in the present invention can

be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, vaginal, or rectal means.

[0237] In addition to the active ingredients (i.e., the HGPRBMY8 nucleic acid or polypeptide, or functional fragments thereof), the pharmaceutical compositions may contain suitable pharmaceutically acceptable carriers or excipients comprising auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration are provided in the latest edition of Remington's Pharmaceutical Sciences (Mack Publishing Co., Easton, PA).

[0238] Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

[0239] Pharmaceutical preparations for oral use can be obtained by the combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropyl-methylcellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth, and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a physiologically acceptable salt thereof, such as sodium alginate.

[0240] Dragee cores may be used in conjunction with physiologically suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or

pigments may be added to the tablets or dragee coatings for product identification, or to characterize the quantity of active compound, i.e., dosage.

[0241] Pharmaceutical preparations, which can be used orally, include push-fit capsules made of gelatin, as well as soft, scaled capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

[0242] Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances, which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. In addition, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyloleate or triglycerides, or liposomes. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

[0243] For topical or nasal administration, penetrants or permeation agents that are appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

[0244] The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

[0245] The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, and the like. Salts tend to be more soluble in aqueous solvents, or other protonic solvents, than are the corresponding free base forms. In

other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, combined with a buffer prior to use. After the pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HGPRBMY8 product, such labeling would include amount, frequency, and method of administration.

[0246] Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose or amount is well within the capability of those skilled in the art. For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., using neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used and extrapolated to determine useful doses and routes for administration in humans.

[0247] A therapeutically effective dose refers to that amount of active ingredient, for example, HGPRBMY8 polypeptide, or fragments thereof, antibodies to HGPRBMY8 polypeptide, agonists, antagonists or inhibitors of HGPRBMY8 polypeptide, which ameliorates, reduces, or eliminates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED_{50} (the dose therapeutically effective in 50% of the population) and LD_{50} (the dose lethal to 50% of the population). The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the ratio, LD_{50}/ED_{50}. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used in determining a range of dosages for human use. Preferred dosage contained in a pharmaceutical composition is within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

[0248] The practitioner, who will consider the factors related to the individual requiring treatment, will determine the exact dosage. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors, which may be taken into account, include the severity of the individual's disease state, general health of the patient, age, weight, and gender of the patient, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/ response to therapy. As a general guide, long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or once every two weeks, depending on half-life and clearance rate of the particular formulation. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

[0249] Normal dosage amounts may vary from 0.1 to 100,000 micrograms (μg), up to a total dose of about 1 gram (g), depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and is generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, and the like.

[0250] In another embodiment of the present invention, antibodies which specifically bind to the HGPRBMY8 polypeptide may be used for the diagnosis of conditions or diseases characterized by expression (or overexpression) of the HGPRBMY8 polynucleotide or polypeptide, or in assays to monitor patients being treated with the HGPRBMY8 polypeptide, or its agonists, antagonists, or inhibitors. The antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for use in therapeutic methods. Diagnostic assays for the HGPRBMY8 polypeptide include methods, which utilize the antibody and a label to detect the protein in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules, which are known in the art, may be used, several of which are described above. In particular, a method of detecting a G-protein coupled receptor, homologue,

or an antibody-reactive fragment thereof, in a sample, comprising: a) contacting the sample with an antibody specific for the polypeptide, or an antigenic fragment thereof, under conditions in which an antigen-antibody complex can form between the antibody and the polypeptide or antigenic fragment thereof in the sample; and b) detecting an antigen-antibody complex formed in step (a), wherein detection of the complex indicates the presence of an antigenic fragment thereof, in the sample.

[0251] The use of mammalian cell reporter assays to demonstrate functional coupling of known GPCRs (G Protein Coupled Receptors) has been well documented in the literature (Gilman, 1987, Boss et al., 1996; Alam & Cook, 1990; George et al., 1997; Selbie & Hill, 1998; Rees et al., 1999). In fact, reporter assays have been successfully used for identifying novel small molecule agonists or antagonists against GPCRs as a class of drug targets (Zlokarnik et al., 1998; George et al., 1997; Boss et al., 1996; Rees et al, 2001). In such reporter assays, a promoter is regulated as a direct consequence of activation of specific signal transduction cascades following agonist binding to a GPCR (Alam & Cook 1990; Selbie & Hill, 1998; Boss et al., 1996; George et al., 1997; Gilman, 1987).

[0252] A number of response element-based reporter systems have been developed that enable the study of GPCR function. These include cAMP response element (CRE)-based reporter genes for G alpha i/o, G alpha s- coupled GPCRs, Nuclear Factor Activator of Transcription (NFAT)-based reporters for G alpha q/11 11or the promiscuous G protein G alpha 15/16 –coupled receptors and MAP kinase reporter genes for use in Galpha i/o coupled receptors (Selbie & Hill, 1998; Boss et al., 1996; George et al., 1997; Blahos, et al., 2001; Offermann & Simon, 1995; Gilman, 1987; Rees et al., 2001). Transcriptional response elements that regulate the expression of Beta-Lactamase within a CHO K1 cell line (CHO-NFAT/CRE: Aurora Biosciences ^{TM}) (Zlokarnik et al., 1998) have been implemented to characterize the function of the orphan HGPRBMY8 polypeptide of the present invention. The system enables demonstration of constitutive G-protein coupling to endogenous cellular signaling components upon intracellular overexpression of orphan receptors. Overexpression has been shown to represent a physiologically relevant event. For example, it has been shown that overexpression occurs in nature during metastatic

carcinomas, wherein defective expression of the monocyte chemotactic protein 1 receptor, CCF2, in macrophages is associated with the incidence of human ovarian carcinoma (Sica, et al.,2000; Salcedo et al., 2000). Indeed, it has been shown that overproduction of the Beta 2 Adrenergic Receptor in transgenic mice leads to constitutive activation of the receptor signaling pathway such that these mice exhibit increased cardiac output (Kypson et al., 1999; Dorn et al., 1999). These are only a few of the many examples demonstrating constitutive activation of GPCRs whereby many of these receptors are likely to be in the active, R*, conformation (J.Wess 1997) (see Example 7).

[0253] Several assay protocols including ELISA, RIA, and FACS for measuring HGPRBMY8 polypeptide are known in the art and provide a basis for diagnosing altered or abnormal levels of HGPRBMY8 polypeptide expression. Normal or standard values for HGPRBMY8 polypeptide expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to the HGPRBMY8 polypeptide under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods; photometric means are preferred. Quantities of HGPRBMY8 polypeptide expressed in subject sample, control sample, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

Microarrays and Screening Assays

[0254] In another embodiment of the present invention, oligonucleotides, or longer fragments derived from the HGPRBMY8 polynucleotide sequence described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously (to produce a transcript image), and to identify genetic variants, mutations and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disease, to diagnose disease, and to develop and monitor the activities of therapeutic agents. In a particular aspect, the microarray is prepared and used

according to the methods described in WO 95/11995 (Chee et al.); D.J. Lockhart et al., 1996, Nature Biotechnology, 14:1675-1680; and M. Schena et al., 1996, Proc. Natl. Acad. Sci. USA, 93:10614-10619). Microarrays are further described in U.S. Patent No. 6,015,702 to P. Lal et al.

[0255] In another embodiment of this invention, the nucleic acid sequence, which encodes the HGPRBMY8 polypeptide, may also be used to generate hybridization probes, which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial PI constructions, or single chromosome cDNA libraries, as reviewed by C.M. Price, 1993, *Blood Rev.*, 7:127-134 and by B.J. Trask, 1991, *Trends Genet.*, 7:149-154.

[0256] Fluorescent *In Situ* Hybridization (FISH), (as described in I. Verma et al., 1988, Human Chromosomes: A Manual of Basic Techniques Pergamon Press, New York, NY) may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in numerous scientific journals, or at Online Mendelian Inheritance in Man (OMIM). Correlation between the location of the gene encoding the HGPRBMY8 polypeptide on a physical chromosomal map and a specific disease, or predisposition to a specific disease, may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences, particularly that of SEQ ID NO:1, or fragments thereof, according to this invention may be used to detect differences in gene sequences between normal, carrier, or affected individuals.

[0257] *In situ* hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers, even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease

or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (R.A. Gatti et al., 1988, Nature, 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the present invention may also be used to detect differences in the chromosomal location due to translocation, inversion, and the like, among normal, carrier, or affected individuals.

[0258] In another embodiment of the present invention, the HGPRBMY8 polypeptide, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between HGPRBMY8 polypeptide, or portion thereof, and the agent being tested, may be measured utilizing techniques commonly practiced in the art. In particular, a method of screening a library of molecules or compounds with an HGPRBMY8 polynucleotide, or fragment thereof, to identify at least one molecule or compound therein which specifically binds to the G-protein coupled receptor polynucleotide sequence, preferably the HGPRBMY8 polynucleotide sequence, or fragment thereof, comprising: a) combining the G-protein coupled receptor polynucleotide, or fragment thereof, with a library of molecules or compounds under conditions to allow specific binding; and b) detecting specific binding, thereby identifying a molecule or compound, which specifically binds to a G-protein coupled receptor-encoding polynucleotide sequence. In a further embodiment, the screening method is a high throughput screening method. Preferably, the library is selected from the group consisting of DNA molecules, RNA molecules, artificial chromosome constructions, PNAs, peptides and proteins. In another preferred embodiment, the candidate small molecules or compounds are a drug or therapeutic.

[0259] In yet another embodiment, a method of screening for candidate compounds capable of modulating activity of a G-protein coupled receptor-encoding polypeptide, comprising: a) contacting a test compound with a cell or tissue expressing the G-protein coupled receptor polypeptide, homologue, or fragment

thereof; and b) selecting as candidate modulating compounds those test compounds that modulate activity of the G-protein coupled receptor polypeptide. Preferably, the candidate compounds are agonists or antagonists of G-protein coupled receptor activity. More preferably, the polypeptide activity is associated with the brain.

[0260] Another technique for drug screening, which may be used, provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in WO 84/03564 (Venton, et al.). In this method, as applied to the HGPRBMY8 protein, large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with the HGPRBMY8 polypeptide, or fragments thereof, and washed. Bound HGPRBMY8 polypeptide is then detected by methods well known in the art. Purified HGPRBMY8 polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

[0261] In a further embodiment of this invention, competitive drug screening assays can be used in which neutralizing antibodies, capable of binding the HGPRBMY8 polypeptide, specifically compete with a test compound for binding to the HGPRBMY8 polypeptide. In this manner, the antibodies can be used to detect the presence of any peptide, which shares one or more antigenic determinants with the HGPRBMY8 polypeptide.

[0262] Other screening and small molecule (e.g., drug) detection assays which involve the detection or identification of small molecules or compounds that can bind to a given protein, i.e., the HGPRBMY8 polypeptide, are encompassed by the present invention. Particularly preferred are assays suitable for high throughput screening methodologies. In such binding-based screening or detection assays, a functional assay is not typically required. All that is needed is a target protein, preferably substantially purified, and a library or panel of compounds (e.g., ligands, drugs, small molecules) to be screened or assayed for binding to the protein target. Preferably, most small molecules that bind to the target protein will modulate activity in some manner, due to preferential, higher affinity binding to functional areas or sites on the

protein.

[0263] An example of such an assay is the fluorescence based thermal shift assay (3-Dimensional Pharmaceuticals, Inc., 3DP; Exton, PA) as described in U.S. Patent Nos. 6,020,141 and 6,036,920 to Pantoliano et al.; see also, J. Zimmerman, 2000, *Gen. Eng. News*, 20(8)). The assay allows the detection of small molecules (e.g., drugs, ligands) that bind to expressed, and preferably purified, HGPRBMY8 polypeptide based on affinity of binding determinations by analyzing thermal unfolding curves of protein-drug or ligand complexes. The drugs or binding molecules determined by this technique can be further assayed, if desired, by methods, such as those described herein, to determine if the molecules affect or modulate function or activity of the target protein.

EXAMPLES

[0264] The Examples herein are meant to exemplify the various aspects of carrying out the invention and are not intended to limit the scope of the invention in any way. The Examples do not include detailed descriptions for conventional methods employed, such as in the construction of vectors, the insertion of cDNA into such vectors, or the introduction of the resulting vectors into the appropriate host. Such methods are well known to those skilled in the art and are described in numerous publications, for example, Sambrook, Fritsch, and Maniatis, Molecular Cloning: a Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, USA, (1989).

EXAMPLE 1
BIOINFORMATICS ANALYSIS

[0265] G-protein coupled receptor sequences were used as a probes to search human genomic sequence databases. The search program used was gapped BLAST (S.F. Altschul, et al., Nuc. Acids Res., 25:3389-4302 (1997)). The top genomic exon hits from the BLAST results were searched back against the non-redundant protein and patent sequence databases. From this analysis, exons encoding potential full-length sequence of a novel human GPCR, HGPRBMY8, was identified directly from the genomic sequence. The full-length clone of this GPCR was experimentally obtained by RT-PCR using the sequence from genomic data. The complete protein

sequence of HGPRBMY8 was analyzed for potential transmembrane domains. TMPRED program (K. Hofmann and W. Stoffel, Biol. Chem., 347:166 (1993) was used for transmembrane prediction. The program predicted seven transmembrane domains and the predicted domains match with the predicted transmembrane domains of related GPCRs at the sequence level. Based on sequence, structure and known GPCR signature sequences, the orphan protein, HGPRBMY8 of the present invention, is a novel human GPCR.

EXAMPLE 2
CLONING OF THE NOVEL HUMAN GPCR HGPRBMY8

[0266] HGPRBMY8 was cloned from a human brain cDNA library (Clontech; Palo Alto, CA) by PCR amplification of the predicted cDNA sequence using sequence specific oligonucleotides. The 5' sense oligonucleotide was as follows:

5'-GGCCGAATTCGCAACCTGTCTCACGCCCTCTGG-3' (SEQ ID NO:5). The 3' anti-sense oligonucleotide was as follows:

5'-GGCCGAATTCGGACAGTTCAAGGTTTGCCTTAGAAC-3' (SEQ ID NO:6). These oligonucleotides contained EcoRI restriction enzyme sites for subcloning the PCR fragment into the mammalian expression vector, pcDNA6. Samples containing human brain cDNA, the 5 prime sense, and 3 prime anti-sense oligonucleotides were subjected to PCR amplification followed by gel purification of the amplified product. The inserts of cDNA clones that were positive by PCR were sized, and two of the largest clones (~1.6 Kb) were sequenced using conventional sequencing methods. Purified sample was digested with EcoRI, extracted with phenol:chloroform, and ligated into pcDNA6. The resultant plasmids were subjected to DNA sequencing and the sequences were verified by comparison with the database sample.

EXAMPLE 3
EXPRESSION PROFILING OF NOVEL HUMAN GPCR, HGPRBMY8

[0267] The oligonucleotides used for the expression profiling of HGPRBMY8 are:

HGPRBMY8-2s: 5'-GCAGAGCACTCCTCCACTCT-3' (SEQ ID NO:34)

HGPRBMY8-2a: 5'-AGCAGGCAATCATGACAATC-3' (SEQ ID NO:35)

[0268] These oligonucleotides were used to measure the steady state levels of mRNA by quantitative PCR. Briefly, first strand cDNA was made from commercially available mRNA (Clontech; Palo Alto, CA). The relative amount of cDNA used in each assay (2.5 ng of cDNA per assay) was determined by performing a parallel experiment using a primer pair for the cyclophilin gene, which is expressed in equal amounts in all tissues. The cyclophilin primer pair detected small variations in the amount of cDNA in each sample, and these data were used for normalization of the data obtained with the primer pair for HGPRBMY8. The PCR data were converted into a relative assessment of the difference in transcript abundance among the tissues tested and the data are presented in Figure 7. Transcripts corresponding to the orphan GPCR, HGPRBMY8, were found to be highly expressed in brain.

EXAMPLE 4
G-PROTEIN COUPLED RECEPTOR PCR EXPRESSION PROFILING

[0269] Based on HGPRBMY8's expression in the brain, further analysis was carried out to determine if there was any additional specificity within sub regions. The same PCR primer pair that was used to identify HGPRBMY8 (also referred to as GPCR 58 and GPCR84) cDNA clones was used to measure the steady state levels of mRNA by quantitative PCR.

GPCR84-s	GTTAGCCTCACCCACCTGTT	(SEQ ID NO:36)
GPCR84-a	CACAATCCAGGTGCCATAGA	(SEQ ID NO:37)

[0270] Briefly, first strand cDNA was made from commercially available brain subregion mRNA (Clontech) and subjected to real time quantitative PCR using a PE 5700 instrument (Applied Biosystems; Foster City, CA) which detects the amount of DNA amplified during each cycle by the fluorescent output of SYBR green, a DNA binding dye specific for double strands. The specificity of the primer pair for its target is verified by performing a thermal denaturation profile at the end of the run which gives an indication of the number of different DNA sequences present by determining melting Tm. In the case of the HGPRBMY8 primer pair, only one DNA fragment was detected having a homogeneous melting point. Contributions of

contaminating genomic DNA to the assessment of tissue abundance is controlled for by performing the PCR with first strand made with and without reverse transcriptase. In all cases, the contribution of material amplified in the no reverse transcriptase controls was negligible.

[0271] More specifically, since HGPRBMY8 is expressed at extremely low levels, each PCR reaction contained the amount of first strand cDNA made from 100 nanograms of poly A+ RNA (2.5 nanograms is the standard amount).

[0272] The number of reactions and amount of mix needed was first determined. All of the samples were run in triplicate, so sample tubes needed 3.5 reactions worth of mixture using the following formula as a guide (2x # tissue samples + 1 no template control + 1 for pipetting error)(3.5).

[0273] The reaction mixture consisted of the following components and volumes:

COMPONENTS	VOL/RXN
2X SybrGreen Master Mix	25 microliters
water	23.5 microliters
primer mix (10uM ea.)	0.5 microliters
cDNA (100ng/uL)	1 microliter

The mixture was initially made without cDNA for enough reactions as determined above. The mix (171.5 µl) was then aliquoted into sample tubes. cDNA (3.5 µl) was added to each sample tube, mixed gently, and spun down for collection. Three 50 µl samples were aliquoted to the optical plate, where the primer and sample were set up for sample analysis. The threshold was set in Log view to intersect linear regions of amplification. The background was set in Linear view to 2-3 cycles before the amplification curve appears. The mean values for RT+ was calculated and normalized to Cyclophilin: dc_t = sample mean – cyclophilin mean. The ddc_t was determined by subtracting individual dcts from the highest value of dc_t in the list. The relative abundance was determined by formula 2^ddc_t.

[0274] Small variations in the amount of cDNA used in each tube was determined by performing a parallel experiment using a primer pair for a gene

expressed in equal amounts in all tissues, cyclophilin. These data were used to normalize the data obtained with the HGPRBMY8 primer pair. The PCR data was converted into a relative assessment of the difference in transcript abundance amongst the tissues tested and the data are presented in bar graph form. Transcripts corresponding to HGPRBMY8 are expressed approximately 825 times greater in the caudate nucleus than in the substantia nigra. Low level expression was detected in the thalamus, amygdala, hippocampus, cerebellum and corpus collosum (see FIG. 8).

EXAMPLE 5
SIGNAL TRANSDUCTION ASSAYS

[0275] The activity of GPCRs or homologues thereof, can be measured using any assay suitable for the measurement of the activity of a G protein-coupled receptor, as commonly known in the art. Signal transduction activity of a G protein-coupled receptor can be monitor by monitoring intracellular Ca^{2+}, cAMP, inositol 1,4,5-triphosphate (IP_3), or 1,2-diacylglycerol (DAG). Assays for the measurement of intracellular Ca^{2+} are described in Sakurai et al. (EP 480 381). Intracellular IP_3 can be measured using a kit available from Amersham, Inc. (Arlington Heights, IL). A kit for measuring intracellular cAMP is available from Diagnostic Products, Inc. (Los Angeles, CA).

[0276] Activation of a G protein-coupled receptor triggers the release of Ca^{2+} ions sequestered in the mitochondria, endoplasmic reticulum, and other cytoplasmic vesicles into the cytoplasm. Fluorescent dyes, e.g., fura-2, can be used to measure the concentration of free cytoplasmic Ca^{2+}. The ester of fura-2, which is lipophilic and can diffuse across the cell membrane, is added to the media of the host cells expressing GPCRs. Once inside the cell, the fura-2 ester is hydrolyzed by cytosolic esterases to its non-lipophilic form, and then the dye cannot diffuse back out of the cell. The non-lipophilic form of fura-2 will fluoresce when it binds to free Ca^{2+}. The fluorescence can be measured without lysing the cells at an excitation spectrum of 340 nm or 380 nm and at fluorescence spectrum of 500 nm (Sakurai et al., EP 480 381).

[0277] Upon activation of a G protein-coupled receptor, the rise of free

cytosolic Ca^{2+} concentrations is preceded by the hydrolysis of phosphatidylinositol 4,5-bisphosphate. Hydrolysis of this phospholipid by the phospholipase C yields 1,2-diacylglycerol (DAG), which remains in the membrane, and water-soluble inositol 1,4,5-triphosphate (IP_3). Binding of ligands or agonists will increase the concentration of DAG and IP3. Thus, signal transduction activity can be measured by monitoring the concentration of these hydrolysis products.

[0278] To measure the IP_3 concentrations, radioactivity labeled ^{3}H-inositol is added to the media of host cells expressing GPCRs. The ^{3}H-inositol is taken up by the cells and incorporated into IP_3. The resulting inositol triphosphate is separated from the mono and di-phosphate forms and measured (Sakurai et al., EP 480 381). Alternatively, Amersham provides an inositol 1,4,5-triphosphate assay system. With this system Amersham provides tritylated inositol 1,4,5-triphosphate and a receptor capable of distinguishing the radioactive inositol from other inositol phosphates. With these reagents an effective and accurate competition assay can be performed to determine the inositol triphosphate levels.

[0279] Cyclic AMP levels can be measured according to the methods described in Gilman et al., Proc. Natl. Acad. Sci. 67:305-312 (1970). In addition, a kit for assaying levels of cAMP is available from Diagnostic Products Corp. (Los Angeles, CA).

EXAMPLE 6
GPCR ACTIVITY

[0280] Another method for screening compounds which are antagonists, and thus inhibit activation of the receptor polypeptide of the present invention is provided. This involves determining inhibition of binding of labeled ligand, such as dATP, dAMP, or UTP, to cells which have the receptor on the surface thereof, or cell membranes containing the receptor. Such a method further involves transfecting a eukaryotic cell with DNA encoding the GPCR polypeptide such that the cell expresses the receptor on its surface. The cell is then contacted with a potential antagonist in the presence of a labeled form of a ligand, such as dATP, dAMP, or UTP. The ligand can be labeled, e.g., by radioactivity, fluorescence, or any detectable

label commonly known in the art. The amount of labeled ligand bound to the receptors is measured, e.g., by measuring radioactivity associated with transfected cells or membrane from these cells. If the compound binds to the receptor, the binding of labeled ligand to the receptor is inhibited as determined by a reduction of labeled ligand which binds to the receptors. This method is called a binding assay. Naturally, this same technique can be used to determine agonists.

[0281] In a further screening procedure, mammalian cells, for example, but not limited to, CHO, HEK 293, Xenopus Oocytes, RBL-2H3, etc., which are transfected, are used to express the receptor of interest. The cells are loaded with an indicator dye that produces a fluorescent signal when bound to calcium, and the cells are contacted with a test substance and a receptor agonist, such as DATP, DAMP, or UTP. Any change in fluorescent signal is measured over a defined period of time using, for example, a fluorescence spectrophotometer or a fluorescence imaging plate reader. A change in the fluorescence signal pattern generated by the ligand indicates that a compound is a potential antagonist or agonist for the receptor.

[0282] In yet another screening procedure, mammalian cells are transfected to express the receptor of interest, and are also transfected with a reporter gene construct that is coupled to activation of the receptor (for example, but not limited to luciferase or beta-galactosidase behind an appropriate promoter). The cells are contacted with a test substance and the receptor agonist (ligand), such as dATP, dAMP, or UTP, and the signal produced by the reporter gene is measured after a defined period of time. The signal can be measured using a luminometer, spectrophotometer, fluorimeter, or other such instrument appropriate for the specific reporter construct used. Inhibition of the signal generated by the ligand indicates that a compound is a potential antagonist for the receptor.

[0283] Another screening technique for antagonists or agonists involves introducing RNA encoding the GPCR polypeptide into cells (or CHO, HEK 293, RBL-2H3, etc.) to transiently or stably express the receptor. The receptor cells are then contacted with the receptor ligand, such as dATP, dAMP, or UTP, and a compound to be screened. Inhibition or activation of the receptor is then determined by detection of a signal, such as, cAMP, calcium, proton, or other ions.

EXAMPLE 7
FUNCTIONAL CHARACTERIZATION OF HGPRBMY8

[0284] The putative GPCR HGPRBMY8 cDNA was PCR amplified using PFU™ (Stratagene). The primers used in the PCR reaction were specific to the HGPRBMY8 polynucleotide and were ordered from Gibco BRL (5 prime primer: 5'-GTCCCCAAGCTTGCACCATGACGTCCACCTGCACCAACAGCA -3' (SEQ ID NO:38). The following 3 prime primer was used to add a Flag-tag epitope to the HGPRBMY8 polypeptide for immunocytochemistry: 5'-CGGGATCCTACTTGTCGTCGTCCTTGTAGTCCATAGGAAAAGTAGCAG AATCGTAGGAA-3' (SEQ ID NO:39). The product from the PCR reaction was isolated from a 0.8% Agarose gel (Invitrogen) and purified using a Gel Extraction Kit ™ from Qiagen.

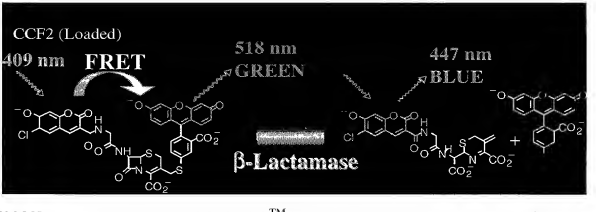
[0285] The purified product was then digested overnight along with the pcDNA3.1 Hygro ™ mammalian expression vector from Invitrogen using the HindIII and BamHI restriction enzymes (New England Biolabs). These digested products were then purified using the Gel Extraction Kit ™ from Qiagen and subsequently ligated to the pcDNA3.1 Hygro ™ expression vector using a DNA molar ratio of 4 parts insert: 1 vector. All DNA modification enzymes were purchased from NEB. The ligation was incubated overnight at 16 degrees Celsius, after which time, one microliter of the mix was used to transform DH5 alpha cloning efficiency competent *E. coli* ™ (Gibco BRL). A detailed description of the pcDNA3.1 Hygro ™ mammalian expression vector is available at the Invitrogen web site (www.Invitrogen.com). The plasmid DNA from the ampicillin resistant clones were isolated using the Wizard DNA Miniprep System ™ from Promega. Positive clones were then confirmed and scaled up for purification using the Qiagen Maxiprep ™ plasmid DNA purification kit.

Cell Line Generation:

[0286] The pcDNA3.1hygro vector containing the orphan HGPRBMY8 cDNA were used to transfect CHO-NFAT/CRE (Aurora Biosciences) cells using Lipofectamine 2000 ™ according to the manufacturers specifications (Gibco BRL).

Two days later, the cells were split 1:3 into selective media (DMEM 11056, 600 µg/ml Hygromycin, 200 µg/ml Zeocin, 10% FBS). All cell culture reagents were purchased from Gibco BRL-Invitrogen.

[0287] The CHO-NFAT/CRE and the CHO-NFAT G alpha 15 cell lines, transiently or stably transfected with the orphan HGPRBMY8 GPCR, were analyzed using the FACS Vantage SE ™ (BD), fluorescence microscopy (Nikon), and the LJL Analyst ™ (Molecular Devices). In this system, changes in real-time gene expression, as a consequence of constitutive G-protein coupling of the orphan HGPRBMY8 GPCR, is examined by analyzing the fluorescence emission of the transformed cells at 447nm and 518nm. The changes in gene expression can be visualized using Beta-Lactamase as a reporter, that, when induced by the appropriate signaling cascade, hydrolyzes an intracellularly loaded, membrane-permeant ester substrate Cephalosporin-Coumarin-Fluorescein2/ Acetoxymethyl (CCF2/AM™ Aurora Biosciences; Zlokarnik, et al., 1998). The CCF2/AM™ substrate is a 7-hydroxycoumarin cephalosporin with a fluorescein attached through a stable thioether linkage. Induced expression of the Beta-Lactamase enzyme is readily apparent since each enzyme molecule produced is capable of changing the fluorescence of many CCF2/AM ™ substrate molecules. A schematic of this cell based system is shown below.



[0288] In summary, CCF2/AM ™ is a membrane permeant, intracellularly-trapped, fluorescent substrate with a cephalosporin core that links a 7-hydroxycoumarin to a fluorescein. For the intact molecule, excitation of the coumarin at 409 nm results in Fluorescence Resonance Energy Transfer (FRET) to the fluorescein which emits green light at 518 nm. Production of active Beta-Lactamase

results in cleavage of the Beta-Lactam ring, leading to disruption of FRET, and excitation of the coumarin only - thus giving rise to blue fluorescent emission at 447 nm.

[0289] Fluorescent emissions were detected using a Nikon-TE300 microscope equipped with an excitation filter (D405/10X-25), dichroic reflector (430DCLP), and a barrier filter for dual DAPI/FITC (510nM) to visually capture changes in Beta-Lactamase expression. The FACS Vantage SE is equiped with a Coherent Enterprise II Argon Laser and a Coherent 302C Krypton laser. In flow cytometry, UV excitation at 351-364 nm from the Argon Laser or violet excitation at 407 nm from the Krypton laser are used. The optical filters on the FACS Vantage SE are HQ460/50m and HQ535/40m bandpass separated by a 490 dichroic mirror.

[0290] Prior to analyzing the fluorescent emissions from the cell lines as described above, the cells were loaded with the CCF2/AM substrate. A 6X CCF2/AM loading buffer was prepared whereby 1mM CCF2/AM (Aurora Biosciences) was dissolved in 100% DMSO (Sigma). Stock solution (12 µl) was added to 60 µl of 100mg/ml Pluronic F127 (Sigma) in DMSO containing 0.1% Acetic Acid (Sigma). This solution was added while vortexing to 1 mL of Sort Buffer (PBS minus calcium and magnesium-Gibco-25 mM HEPES-Gibco- pH 7.4, 0.1% BSA). Cells were placed in serum-free media and the 6X CCF2/AM was added to a final concentration of 1X. The cells were then loaded at room temperature for one to two hours, and then subjected to fluorescent emission analysis as described herein. Additional details relative to the cell loading methods and/or instrument settings may be found by reference to the following publications: see Zlokarnik, et al., 1998; Whitney et al., 1998; and BD Biosciences,1999.

Immunocytochemistry:

[0291] The cell lines transfected and selected for expression of Flag-epitope tagged orphan GPCRs were analyzed by immunocytochemistry. The cells were plated at $1X10^3$ in each well of a glass slide (VWR). The cells were rinsed with PBS followed by acid fixation for 30 minutes at room temperature using a mixture of 5% Glacial Acetic Acid / 90% ethanol. The cells were then blocked in 2% BSA and

0.1%Triton in PBS, incubated for 2 h at room temperature or overnight at 4°C. A monoclonal FITC antibody directed against FLAG was diluted at 1:50 in blocking solution and incubated with the cells for 2 h at room temperature. Cells were then washed three times with 0.1%Triton in PBS for five minutes. The slides were overlayed with mounting media dropwise with Biomedia –Gel Mount™ (Biomedia; Containing Anti-Quenching Agent). Cells were examined at 10x magnification using the Nikon TE300 equiped with FITC filter (535nm).

[0292] There is strong evidence that certain GPCRs exhibit a cDNA concentration-dependent constitutive activity through cAMP response element (CRE) luciferase reporters (Chen et al., 1999). In an effort to demonstrate functional coupling of HGPRBMY8 to known GPCR second messenger pathways, the HGPRBMY8 polypeptide was expressed at high constitutive levels in the CHO-NFAT/CRE cell line. To this end, the HGPRBMY8 cDNA was PCR amplified and subcloned into the pcDNA3.1 hygro ™ mammalian expression vector as described herein. Early passage CHO-NFAT/CRE cells were then transfected with the resulting pcDNA3.1 hygro ™ / HGPRBMY8 construct. Transfected and non-transfected CHO-NFAT/CRE cells (control) were loaded with the CCF2 substrate and stimulated with 10 nM PMA, 1 μM Thapsigargin (NFAT stimulator), and 10 μM Forskolin (CRE stimulator) to fully activate the NFAT/CRE element. The cells were then analyzed for fluorescent emission by FACS.

[0293] The FACS profile demonstrates the constitutive activity of HGPRBMY8 in the CHO-NFAT/CRE line as evidenced by the significant population of cells with blue fluorescent emission at 447 nm (see Figure 12: Blue Cells). Figure 12 further describes CHO-NFAT/CRE cell lines transfected with the pcDNA3.1 Hygro ™ / HGPRBMY8 mammalian expression vector. The cells were analyzed via FACS according to their wavelength emission at 518 nM (Channel R3 - Green Cells), and 447 nM (Channel R2 – Blue Cells). As shown, overexpression of HGPRBMY8 results in functional coupling and subsequent activation of beta lactamase gene expression, as evidenced by the significant number of cells with fluorescent emission at 447 nM relative to the non-transfected control CHO-NFAT/CRE cells (shown in Figure 11). As expected, the NFAT/CRE response element in the untransfected

control cell line was not activated (i.e., beta lactamase not induced), enabling the CCF2 substrate to remain intact, and resulting in the green fluorescent emission at 518 nM (see Figure 11-Green Cells). Figure 11 describes control CHO-NFAT/CRE (Nuclear Factor Activator of Transcription (NFAT) / cAMP response element (CRE)) cell lines, in the absence of the pcDNA3.1 Hygro ^{TM} / HGPRBMY8 mammalian expression vector transfection. The cells were analyzed via FACS (Fluorescent Assisted Cell Sorter) according to their wavelength emission at 518 nM (Channel R3 - Green Cells), and 447 nM (Channel R2 – Blue Cells). As shown, the vast majority of cells emit at 518 nM, with minimal emission observed at 447 nM. The latter is expected since the NFAT/CRE response elements remain dormant in the absence of an activated G-protein dependent signal transduction pathway (e.g., pathways mediated by Gq/11 or Gs coupled receptors). As a result, the cell permeant, CCF2/AM^{TM} (Aurora Biosciences; Zlokarnik, et al., 1998) substrate remains intact and emits light at 518 nM.

[0294] A very low level of leaky Beta Lactamase expression was detectable as evidenced by the small population of cells emitting at 447 nm. Analysis of a stable pool of cells transfected with HGPRBMY8 revealed constitutive coupling of the cell population to the NFAT/CRE response element, activation of Beta Lactamase and cleavage of the substrate (Figure 12-Blue Cells). These results demonstrate that overexpression of HGPRBMY8 leads to constitutive coupling of signaling pathways known to be mediated by Gq/11or G alpha 15/16 or Gs coupled receptors that converge to activate either the NFAT or CRE response elements respectively (Boss et al., 1996; Chen et al., 1999).

[0295] In an effort to further characterize the observed functional coupling of the HGPRBMY8 polypeptide, its ability to couple to the cAMP response element (CRE) independent of the NFAT response element was examined. To this end, HEK-CRE cell line that contained only the integrated 3XCRE linked to the Beta-Lactamase reporter was transfected with the pcDNA3.1 hygro ^{TM} / HGPRBMY8 construct. Analysis of the fluorescence emission from this stable pool showed that HGPRBMY8 constitutively coupled to the cAMP mediated second messenger pathways (see Figure 14 relative to Figure 13). Figure 14 describes FACS analysis of HEK-CRE cell lines

transfected with the pcDNA3.1 Hygro ™ / HGPRBMY8 mammalian expression vector according to their wavelength emission at 518 nM (Channel R3 - Green Cells), and 447 nM (Channel R2 – Blue Cells). As shown, overexpression of HGPRBMY8 in the HEK-CRE cells resulted in functional coupling, as evidenced by the insignificant background level of cells with fluorescent emission at 447 nM. Figure 13 describes HEK-CRE cell lines in the absence of the pcDNA3.1 Hygro ™ / HGPRBMY8 mammalian expression vector transfection. The cells were analyzed via FACS (Fluorescent Assisted Cell Sorter) according to their wavelength emission at 518 nM (Channel R3 - Green Cells), and 447 nM (Channel R2 – Blue Cells). As shown, the vast majority of cells emit at 518 nM, with minimal emission observed at 447 nM. The latter is expected since the CRE response elements remain dormant in the absence of an activated G-protein dependent signal transduction pathway (e.g., pathways mediated by Gs coupled receptors). As a result, the cell permeant, CCF2/AM™ (Aurora Biosciences; Zlokarnik, et al., 1998) substrate remains intact and emits light at 518 nM.

[0296] Experiments have shown that known G coupled receptors do demonstrate constitutive activation when overexpressed in the HEK-CRE cell line. For example, direct activation of adenylate cyclase using 10 μM Forskolin has been shown to activate CRE and the subsequent induction of Beta-Lactamase in the HEK-CRE cell line (data not shown). In conclusion, the results are consistent with HGPRBMY8 representing a functional GPCR analogous to known Gs coupled receptors (Boss et al., 1996).

Demonstration of Cellular Expression:

[0297] HGPRBMY8 was tagged at the C-terminus using the Flag epitope and inserted into the pcDNA3.1 hygro ™ expression vector, as described herein. Immunocytochemistry of CHO-NFAT/CRE cell lines transfected with the Flag-tagged HGPRBMY8 construct with FITC conjugated Anti Flag monoclonal antibody demonstrated that HGPRBMY8 is indeed a cell surface receptor. The immunocytochemistry also confirmed expression of the HGPRBMY8 in the CHO-NFAT/CRE cell lines. Briefly, CHO-NFAT/CRE cell lines were transfected with

pcDNA3.1 hygro ™ / HGPRBMY8-Flag vector, fixed with 70% methanol, and permeablized with 0.1%TritonX100. The cells were then blocked with 1% Serum and incubated with a FITC conjugated Anti Flag monoclonal antibody at 1:50 dilution in PBS-Triton. The cells were then washed several times with PBS-Triton, overlayed with mounting solution, and fluorescent images were captured (see Figure 15A-D). Figure 15 describes CHO-NFAT/CRE cell lines transfected with the pcDNA3.1 Hygro ™ / HGPRBMY8-FLAG mammalian expression vector subjected to immunocytochemistry using an FITC conjugated Anti Flag monoclonal antibody. Panel A shows the transfected CHO-NFAT/CRE cells under visual wavelengths, and panel B shows the fluorescent emission of the same cells at 530 nm after illumination with a mercury light source. The cell expression is clearly evident in panel B, and is consistent with the HGPRBMY8 polypeptide representing a member of the GPCR family. The control cell line, non-transfected CHO-NFAT /CREcell line, exhibited no detectable background fluorescence (Figure 15). The HGPRBMY8 –FLAG tagged expressing CHO-NFAT /CRE line exhibited specific plasma membrane expression as indicated (Figure 15). These data provide clear evidence that HGPRBMY8 is expressed in these cells and the majority of the protein is localized to the cell surface. Cell surface localization in consistent with HGPRBM8 representing a 7 transmembrane domain containing GPCR. Taken together, the data indicate that HGPRBMY8 is a cell surface GPCR that can function through increases in either cAMP or Ca^{2+} signal transduction pathways via G alpha 15.

Screening Paradigm

[0298] The Aurora Beta-Lactamase technology provides a clear path for identifying agonists and antagonists of the HGPRBMY8 polypeptide. Cell lines that exhibit a range of constitutive coupling activity have been identified by sorting through HGPRBMY8 transfected cell lines using the FACS Vantage SE (see Figure 16). For example, cell lines have been sorted that have an intermediate level of orphan GPCR expression, which also correlates with an intermediate coupling response, using the LJL analyst. Such cell lines will provide the opportunity to screen, indirectly, for both agonists and antogonists of HGPRBMY8 by looking for

inhibitors that block the beta lactamase response, or agonists that increase the beta lactamase response. As described herein, modulating the expression level of beta lactamase directly correlates with the level of cleaved CCF2 substrate. For example, this screening paradigm has been shown to work for the identification of modulators of a known GPCR, 5HT6, that couples through Adenylate Cyclase, in addition to, the identification of modulators of the 5HT2c GPCR, that couples through changes in [Ca^{2+}]i. The data shown below represent cell lines that have been engineered with the desired pattern of HGPRBMY8 expression to enable the identification of potent small molecule agonists and antagonists. HGPRBMY8 modulator screens may be carried out using a variety of high throughput methods known in the art, though preferably using the fully automated Aurora UHTSS system. The untransfected CHO-NFAT/CRE cell line represents the relative background level of beta lactamase expression (Figure 16; panel a). Figure 16 describes several CHO-NFAT/CRE cell lines transfected with the pcDNA3.1 Hygro ™ / HGPRBMY8 mammalian expression vector isolated via FACS that had either intermediate or high beta lactamase expression levels of constitutive activation. Panel A shows untransfected CHO-NFAT/CRE cells prior to stimulation with 10 nM PMA, 1 μM Thapsigargin, and 10 μM Forskolin (- P/T/F). Panel B shows CHO-NFAT/CRE cells after stimulation with 10 nM PMA, 1 μM Thapsigargin, and 10 μM Forskolin (+ P/T/F). Panel C shows a representative orphan GPCR (oGPCR) transfected CHO-NFAT/CRE cells that have an intermediate level of beta lactamase expression. Panel D shows a representative orphan GPCR transfected CHO-NFAT/CRE that have a high level of beta lactamase expression. Following treatment with a cocktail of 10 nM PMA, 1μM Thapsigargin, and 10μM Forskolin (Figure 16; P/T/F; panel b), the cells fully activate the CRE-NFAT response element demonstrating the dynamic range of the assay. Panel C (Figure 16) represents an orphan transfected CHO-NFAT/CRE cell line that shows an intermediate level of beta lactamase expression post P/T/F stimulation, while panel D (Figure 16) represents a orphan transfected CHO-NFAT/CRE cell line that shows a high level of constitutive beta lactamase expression.

EXAMPLE 8

G-PROTEIN COUPLED RECEPTOR PCR EXPRESSION PROFILING

[0299] RNA quantification was performed using the Taqman real-time-PCR fluorogenic assay. The Taqman assay is one of the most precise methods for assaying the concentration of nucleic acid templates .

[0300] All cell lines were grown using standard conditions: RPMI 1640 supplemented with 10% fetal bovine serum, 100 IU/ml penicillin, 100 mg/ml streptomycin, and 2 mM L-glutamine, 10 mM Hepes (all from GibcoBRL; Rockville, MD). Eighty percent confluent cells were washed twice with phosphate-buffered saline (GibcoBRL) and harvested using 0.25% trypsin (GibcoBRL). RNA was prepared using the RNeasy Maxi Kit from Qiagen (Valencia, CA).

[0301] cDNA template for real-time PCR was generated using the Superscript First Strand Synthesis system for RT-PCR.

[0302] SYBR Green real-time PCR reactions were prepared as follows: The reaction mix consisted of 20 ng first strand cDNA; 50 nM Forward Primer; 50 nM Reverse Primer; 0.75X SYBR Green I (Sigma); 1X SYBR Green PCR Buffer (50mMTris-HCl pH8.3, 75mM KCl); 10%DMSO; 3mM $MgCl_2$; 300 M each dATP, dGTP, dTTP, dCTP; 1 U Platinum Taq DNA Polymerase High Fidelity (Cat# 11304-029; Life Technologies; Rockville, MD); 1:50 dilution; ROX (Life Technologies). Real-time PCR was performed using an Applied Biosystems 5700 Sequence Detection System. Conditions were 95C for 10 min (denaturation and activation of Platinum Taq DNA Polymerase), 40 cycles of PCR (95C for 15 sec, 60C for 1 min). PCR products are analyzed for uniform melting using an analysis algorithm built into the 5700 Sequence Detection System.

Forward primer: 745 GPCR84-2s: 5'-GCAGAGCACTCCTCCACTCT-3' (SEQ ID NO:34); and

Reverse primer: 746 GPCR84-2a: 5'-AGCAGGCAATCATGACAATC-3' (SEQ ID NO:35).

[0303] cDNA quantification used in the normalization of template quantity was performed using Taqman technology. Taqman reactions are prepared as follows: The reaction mix consisted of 20 ng first strand cDNA; 25 nM GAPDH-F3, Forward Primer; 250 nM GAPDH-R1 Reverse Primer; 200 nM GAPDH-PVIC Taqman Probe (fluorescent dye labeled oligonucleotide primer); 1X Buffer A (Applied Biosystems);

5.5 mM MgCl2; 300 M dATP, dGTP, dTTP, dCTP; 1 U Amplitaq Gold (Applied Biosystems). GAPDH, D-glyceraldehyde -3-phosphate dehydrogenase, was used as control to normalize mRNA levels.

[0304] Real-time PCR was performed using an Applied Biosystems 7700 Sequence Detection System. Conditions were 95C for 10 min. (denaturation and activation of Amplitaq Gold), 40 cycles of PCR (95C for 15 sec, 60C for 1 min).

[0305] The sequences for the GAPDH oligonucleotides used in the Taqman reactions are as follows:

GAPDH-F3 -5'-AGCCGAGCCACATCGCT-3' (SEQ ID NO:60)

GAPDH-R1 -5'-GTGACCAGGCGCCCAATAC-3' (SEQ ID NO:61)

GAPDH-PVIC Taqman® Probe -VIC-5' -
CAAATCCGTTGACTCCGACCTTCACCTT-3' TAMRA (SEQ ID NO:62).

[0306] The Sequence Detection System generates a Ct (threshold cycle) value that is used to calculate a concentration for each input cDNA template. cDNA levels for each gene of interest are normalized to GAPDH cDNA levels to compensate for variations in total cDNA quantity in the input sample. This is done by generating GAPDH Ct values for each cell line. Ct values for the gene of interest and GAPDH are inserted into a modified version of the Ct equation (Applied Biosystems Prism 7700 Sequence Detection System User Bulletin #2), which is used to calculate a GAPDH normalized relative cDNA level for each specific cDNA. The Ct equation is as follows: relative quantity of nucleic acid template $=2^{Ct} = 2^{(Cta-Ctb)}$, where Cta=Ct target – Ct GAPDH, and Ctb = Ct reference – Ct GAPDH. (No reference cell line was used for the calculation of relative quantity; Ctb was defined as 21).

[0307] The Graph # of Table 1 corresponds to the tissue type position number of Figure 17. HGPRBMY8 (also known as GPCR84 or GPCR58) was found to have relatively low expression in the tumor cell lines assayed in the OCLP-1 (oncology cell line panel). HGPRBMY8 message appears to be especially scarce in breast tumor cell lines. The average HGPRBMY8 message message level in lung tumor cell lines is 2-3 fold lower than the average for other cell lines assayed.

TABLE 1

Graph #	Name	Tissue	Ct GAPDH	Ct GPCR84	dCt	ddCt	Quant.
1	AIN 4	breast	17.49	40	22.51	1.51	0.0E+00
2	AIN 4T	breast	17.15	36.8	19.65	-1.35	2.5E+00
3	AIN4/myc	breast	17.81	40	22.19	1.19	0.0E+00
4	BT-20	breast	17.9	36.15	18.25	-2.75	6.7E+00
5	BT-474	breast	17.65	38.34	20.69	-0.31	1.2E+00
6	BT-483	breast	17.45	35.6	18.15	-2.85	7.2E+00
7	BT-549	breast	17.55	38.21	20.66	-0.34	1.3E+00
8	DU4475	breast	18.1	40	21.9	0.9	0.0E+00
9	H3396	breast	18.04	36.71	18.67	-2.33	5.0E+00
10	HBL100	breast	17.02	37.16	20.14	-0.86	1.8E+00
11	Her2 MCF-7	breast	19.26	35.62	16.36	-4.64	2.5E+01
12	HS 578T	breast	17.83	37.28	19.45	-1.55	2.9E+00
13	MCF7	breast	17.83	40	22.17	1.17	0.0E+00
14	MCF-7/AdrR	breast	17.23	36.01	18.78	-2.22	4.7E+00
15	MDAH 2774	breast	16.87	35.24	18.37	-2.63	6.2E+00
16	MDA-MB-175-VII	breast	15.72	34.08	18.36	-2.64	6.2E+00
17	MDA-MB-231	breast	17.62	40	22.38	1.38	0.0E+00
18	MDA-MB-453	breast	17.9	37.57	19.67	-1.33	2.5E+00
19	MDA-MB-468	breast	17.49	37.58	20.09	-0.91	1.9E+00
20	Pat-21 R60	breast	35.59	40	4.41	-16.59	ND
21	SKBR3	breast	17.12	40	22.88	1.88	0.0E+00
22	T47D	breast	18.86	40	21.14	0.14	0.0E+00
23	UACC-812	breast	17.06	38.26	21.2	0.2	8.7E-01

TABLE 1

Graph #	Name	Tissue	Ct GAPDH	Ct GPCR84	dCt	ddCt	Quant.
24	ZR-75-1	breast	15.95	35.36	19.41	-1.59	3.0E+00
25	C-33A	cervical	17.49	36.96	19.47	-1.53	2.9E+00
26	Ca Ski	cervical	17.38	37.78	20.4	-0.6	1.5E+00
27	HeLa	cervical	17.59	40	22.41	1.41	0.0E+00
28	HT-3	cervical	17.42	35.69	18.27	-2.73	6.6E+00
29	ME-180	cervical	16.86	34.57	17.71	-3.29	9.8E+00
30	SiHa	cervical	18.07	40	21.93	0.93	0.0E+00
31	SW756	cervical	15.59	36.45	20.86	-0.14	1.1E+00
32	CACO-2	colon	17.56	40	22.44	1.44	0.0E+00
33	CCD-112Co	colon	18.03	40	21.97	0.97	0.0E+00
34	CCD-33Co	colon	17.07	39.44	22.37	1.37	3.9E-01
35	Colo 205	colon	18.02	40	21.98	0.98	0.0E+00
36	Colo 320DM	colon	17.01	40	22.99	1.99	0.0E+00
37	Colo201	colon	17.89	34.47	16.58	-4.42	2.1E+01
38	Cx-1	colon	18.79	40	21.21	0.21	0.0E+00
39	ddH2O	colon	40	40	0	-21	ND
40	HCT116	colon	17.59	36.22	18.63	-2.37	5.2E+00
41	HCT116/epo5	colon	17.71	36.42	18.71	-2.29	4.9E+00
42	HCT116/ras	colon	17.18	40	22.82	1.82	0.0E+00
43	HCT116/TX15 CR	colon	17.36	36.91	19.55	-1.45	2.7E+00
44	HCT116/vivo	colon	17.7	37.01	19.31	-1.69	3.2E+00
45	HCT116/VM46	colon	17.87	37.55	19.68	-1.32	2.5E+00
46	HCT116/VP35	colon	17.3	40	22.7	1.7	0.0E+00
47	HCT-8	colon	17.44	36.86	19.42	-1.58	3.0E+00

TABLE 1

Graph #	Name	Tissue	Ct GAPDH	Ct GPCR84	dCt	ddCt	Quant.
48	HT-29	colon	17.9	40	22.1	1.1	0.0E+00
49	LoVo	colon	17.64	40	22.36	1.36	0.0E+00
50	LS 174T	colon	17.93	40	22.07	1.07	0.0E+00
51	LS123	colon	17.65	36.05	18.4	-2.6	6.1E+00
52	MIP	colon	16.92	35.65	18.73	-2.27	4.8E+00
53	SK-CO-1	colon	17.75	39.84	22.09	1.09	4.7E-01
54	SW1417	colon	17.22	39.11	21.89	0.89	5.4E-01
55	SW403	colon	18.39	40	21.61	0.61	0.0E+00
56	SW480	colon	17	40	23	2	0.0E+00
57	SW620	colon	17.16	40	22.84	1.84	0.0E+00
58	SW837	colon	18.35	37.65	19.3	-1.7	3.2E+00
59	T84	colon	16.44	40	23.56	2.56	0.0E+00
60	CCD-18Co	colon, fibroblast	17.19	40	22.81	1.81	0.0E+00
61	HT-1080	fibrosarcoma	17.16	40	22.84	1.84	0.0E+00
62	CCRF-CEM	leukemia	17.07	40	22.93	1.93	0.0E+00
63	HL-60	leukemia	17.54	40	22.46	1.46	0.0E+00
64	K562	leukemia	18.42	40	21.58	0.58	0.0E+00
65	A-427	lung	18	40	22	1	0.0E+00
66	A549	lung	17.63	37.06	19.43	-1.57	3.0E+00
67	Calu-3	lung	18.09	37.38	19.29	-1.71	3.3E+00
68	Calu-6	lung	16.62	40	23.38	2.38	0.0E+00
69	ChaGo-K-1	lung	17.79	37.16	19.37	-1.63	3.1E+00
70	DMS 114	lung	18.14	40	21.86	0.86	0.0E+00
71	LX-1	lung	18.17	40	21.83	0.83	0.0E+00

TABLE 1

Graph #	Name	Tissue	Ct GAPDH	Ct GPCR84	dCt	ddCt	Quant.
72	MRC-5	lung	17.3	40	22.7	1.7	0.0E+00
73	MSTO-211H	lung	16.81	40	23.19	2.19	0.0E+00
74	NCI-H596	lung	17.73	40	22.27	1.27	0.0E+00
75	SHP-77	lung	18.66	40	21.34	0.34	0.0E+00
76	Sk-LU-1	lung	15.81	35.83	20.02	-0.98	2.0E+00
77	SK-MES-1	lung	17.1	36.33	19.23	-1.77	3.4E+00
78	SW1271	lung	16.45	40	23.55	2.55	0.0E+00
79	SW1573	lung	17.14	40	22.86	1.86	0.0E+00
80	SW900	lung	18.17	40	21.83	0.83	0.0E+00
81	Hs 294T	melanoma	17.73	35.38	17.65	-3.35	1.0E+01
82	A2780/DDP-R	ovarian	21.51	40	18.49	-2.51	0.0E+00
83	A2780/DDP-S	ovarian	17.89	35.73	17.84	-3.16	8.9E+00
84	A2780/epo5	ovarian	17.54	35.12	17.58	-3.42	1.1E+01
85	A2780/TAX-R	ovarian	18.4	38.33	19.93	-1.07	2.1E+00
86	A2780/TAX-S	ovarian	17.83	40	22.17	1.17	0.0E+00
87	Caov-3	ovarian	15.5	35.35	19.85	-1.15	2.2E+00
88	ES-2	ovarian	17.22	40	22.78	1.78	0.0E+00
89	HOC-76	ovarian	34.3	40	5.7	-15.3	ND
90	OVCAR-3	ovarian	17.09	36.66	19.57	-1.43	2.7E+00
91	PA-1	ovarian	17.33	40	22.67	1.67	0.0E+00
92	SW 626	ovarian	16.94	40	23.06	2.06	0.0E+00
93	UPN251	ovarian	17.69	36.75	19.06	-1.94	3.8E+00
94	LNCAP	prostate	18.17	40	21.83	0.83	0.0E+00
95	PC-3	prostate	17.25	40	22.75	1.75	0.0E+00
96	A431	squamous	19.85	40	20.15	-0.85	0.0E+00

EXAMPLE 9

PHAGE DISPLAY METHODS FOR IDENTIFYING PEPTIDE LIGANDS OR MODULATORS OF ORPHAN GPCRS

Library Construction

[0308] Two HGPRBMY libraries were used for identifying peptides that may function as modulators. Specifically, a 15-mer library was used to identify peptides that may function as agonists or antagonists. The 15-mer library is an aliquot of the 15-mer library originally constructed by G.P. Smith (Scott, JK and Smith, GP. 1990, Science 249:386-390). A 40-mer library was used for identifying natural ligands and constructed essentially as previously described, using an M13 phage library displaying random 38-amino acid peptides as a source of novel sequences with affinity to selected targets (BK Kay, et al. 1993, Gene 128:59-65). This method for constructing the 40-mer library was followed with the exception that a 15 base pair complementary region was used to anneal the two oligonucleotides, as opposed to 6, 9, or 12 base pairs, as described below.

[0309] The oligos used are: Oligo 1: 5'-CGAAGCGTAAGGGCCCAGCCGGCCNN (BNNx19) BCCGGGTCCGGGCGGC - 3' (SEQ ID NO:63) and Oligo2: 5'- AAAAGGAAAAAAGCGGCCGC (VNNx20) GCCGCCCGGACCCGG-3' (SEQ ID NO:64), where N= A+G+C+T and B = C+G+T and V=C+A+G.

[0310] The oligos were are annealed via their 15 base pair complimentary sequences which encode a constant ProGlyProGlyGly (SEQ ID NO:65) pentapeptide sequence between the random 20 amino acid segments, and then extended by standard procedure using Klenow enzyme. This was followed by endonuclease digestion using Sfi1 and Not1 enzymes and ligation to Sfi1 and Not1 cleaved pCantab5E (Pharmacia). The ligation mixture was electroporated into *E. coli* XL1Blue and phage clones were essentially generated as suggested by the manufacturer (Pharmacia) for making ScFv antibody libraries in pCantab5E.

Sequencing Bound Phage

[0311] Standard procedures commonly known in the art were used. Phage in eluates were infected into *E. coli* host strain (TG1 for the 15-mer library; XL1Blue for the 40-mer library) and plated for single colonies. Colonies were grown in liquid and sequenced by standard procedure which involved: 1) generating PCR product with suitable primers of the library segments in the phage genome (15-mer library) or pCantab5E (40-mer library); and 2) sequencing PCR products using one primer of each PCR primer pair. Sequences were visually inspected or were inspected by using the Vector NTI alignment tool.

Peptide Modulators Of The Present Invention

[0312] The following serve as non-limiting examples of HGPRBMY8 peptide modulators:

GDFWYEACESSCAFW	(SEQ ID NO:66)
LEWGSDVFYDVYDCC	(SEQ ID NO:67)
CLRSGTGCAFQLYRF	(SEQ ID NO:68)
NNFPCLRSGRNCDAG	(SEQ ID NO:69)
RIVPNGYFNVHGRSL	(SEQ ID NO:70)
RIDSCAKYFLRSCD	(SEQ ID NO:71)

Peptide Synthesis

[0313] Peptides were synthesized on Fmoc-Knorr amide resin [N-(9-fluorenyl)methoxycarbonyl-Knorr amide-resin, Midwest Biotech, Fishers, IN] with an Applied Biosystems (Foster City, CA) model 433A synthesizer and the *FastMoc* chemistry protocol (0.25mmol scale) supplied with the instrument. Amino acids were double coupled as their N-alpha-Fmoc- derivatives and reactive side chains were protected as follows: Asp, Glu: t-Butyl ester (OtBu); Ser, Thr, Tyr: t-Butyl ether (tBu); Asn, Cys, Gln, His: Triphenylmethyl (Trt); Lys, Trp: t-Butyloxycarbonyl (Boc); Arg: 2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl (Pbf). After the final double coupling cycle, the N-terminal Fmoc group was removed by the multi-step treatment with piperidine in N-Methylpyrrolidone as described by the manufacturer. The N-terminal free amines were then treated with 10% acetic anhydride, 5% Diisopropylamine in N-Methylpyrrolidone to yield the N-acetyl-derivative. The protected peptidyl-resins were simultaneously deprotected and removed from the

resin by standard methods. The lyophilized peptides were purified on C_{18} to apparent homogeneity as judged by RP-HPLC analysis. Predicted peptide molecular weights were verified by electrospray mass spectrometry (J. Biol. Chem. vol. 273, pp.12041-12046, 1998).

[0314] Cyclic analogs were prepared from the crude linear products. The cystine disulfide was formed using one of the following methods:

Method 1:

[0315] A sample of the crude peptide was dissolved in water at a concentration of 0.5 mg/mL and the pH adjusted to 8.5 with NH_4OH. The reaction was stirred at room temperature, and monitored by RP-HPLC. Once complete, the reaction was brought to pH 4 with acetic acid and lyophilized. The product was purified and characterized as above.

Method 2

[0316] A sample of the crude peptide was dissolved at a concentration of 0.5mg/mL in 5% acetic acid. The pH was adjusted to 6.0 with NH_4OH. DMSO (20% by volume) was added and the reaction was stirred overnight. After analytical RP-HPLC analysis, the reaction was diluted with water and triple lyophilized to remove DMSO. The crude product was purified by preparative RP-HPLC. (JACS. vol. 113, 6657, 1991).

Assessing Affect of Peptides on GPCR Function

[0317] The effect of any one of these peptides on the function of the GPCR of the present invention may be determined by adding an effective amount of each peptide to each functional assay. Representative functional assays are described more specifically herein, particularly Example 7.

Uses Of The Peptide Modulators Of The Present Invention

[0318] The aforementioned peptides of the present invention are useful for a variety of purposes, though most notably for modulating the function of the GPCR of the present invention, and potentially with other GPCRs of the same G-protein coupled receptor subclass (e.g., peptide receptors, adrenergic receptors, purinergic receptors, etc.), and/or other subclasses known in the art. For example, the peptide modulators of the present invention may be useful as HGPRBMY8 agonists.

Alternatively, the peptide modulators of the present invention may be useful as HGPRBMY8 antagonists of the present invention. In addition, the peptide modulators of the present invention may be useful as competitive inhibitors of the HGPRBMY8 cognate ligand(s), or may be useful as non-competitive inhibitors of the HGPRBMY8 cognate ligand(s).

[0319] Furthermore, the peptide modulators of the present invention may be useful in assays designed to either deorphan the HGPRBMY8 polypeptide of the present invention, or to identify other agonists or antagonists of the HGPRBMY8 polypeptide of the present invention, particularly small molecule modulators.

EXAMPLE 10

METHOD OF CREATING N- AND C- TERMINAL DELETION MUTANTS CORRESPONDING TO THE HGPRBMY8 POLYPEPTIDE

[0320] As described elsewhere herein, the present invention encompasses the creation of N- and C-terminal deletion mutants, in addition to any combination of N- and C-terminal deletions thereof, corresponding to the HGPRBMY8 polypeptide of the present invention. A number of methods are available to one skilled in the art for creating such mutants. Such methods may include a combination of PCR amplification and gene cloning methodology. Although one of skill in the art of molecular biology, through the use of the teachings provided or referenced herein, and/or otherwise known in the art as standard methods, could readily create each deletion mutants of the present invention, exemplary methods are described below.

[0321] Briefly, using the isolated cDNA clone encoding the full-length HGPRBMY8 polypeptide sequence, appropriate primers of about 15-25 nucleotides derived from the desired 5' and 3' positions of SEQ ID NO:1 may be designed to PCR amplify, and subsequently clone, the intended N- and/or C-terminal deletion mutant. Such primers could comprise, for example, an inititation and stop codon for the 5' and 3' primer, respectively. Such primers may also comprise restriction sites to facilitate cloning of the deletion mutant post amplification. Moreover, the primers may comprise additional sequences, such as, for example, flag-tag sequences, kozac sequences, or other sequences discussed and/or referenced herein.

[0322] For example, in the case of the T36 to P508 N-terminal deletion mutant, the following primers could be used to amplify a cDNA fragment corresponding to this deletion mutant:

5' Primer 5'-GCAGCA GCGGCCGC ACCGTGCTGGTTATCTTCCTCGCCG –3' (SEQ ID NO:72)
NotI
3' Primer 5'- GCAGCA GTCGAC AGGAAAAGTAGCAGAATCGTAGG -3' (SEQ ID NO:73)
SalI

[0323] For example, in the case of the M1 to Y454 C-terminal deletion mutant, the following primers could be used to amplify a cDNA fragment corresponding to this deletion mutant:

5' Primer 5'- GCAGCA GCGGCCGC ATGACGTCCACCTGCACCAACAGC –3' (SEQ ID NO:74)
NotI
3' Primer 5'- GCAGCA GTCGAC ATAGACATAGGGGTGGATGCAGCAC -3' (SEQ ID NO:75)
SalI

[0324] Representative PCR amplification conditions are provided below, although the skilled artisan would appreciate that other conditions may be required for efficient amplification. A 100 µl PCR reaction mixture may be prepared using 10ng of the template DNA (cDNA clone of HGPRBMY8), 200 µM 4dNTPs, 1µM primers, 0.25U Taq DNA polymerase (PE), and standard Taq DNA polymerase buffer. Typical PCR cycling condition are as follows:

20-25 cycles: 45 sec, 93 degrees
2 min, 50 degrees
2 min, 72 degrees
1 cycle: 10 min, 72 degrees

[0325] After the final extension step of PCR, 5U Klenow Fragment may be added and incubated for 15 min at 30 degrees.

[0326] Upon digestion of the fragment with the NotI and SalI restriction

enzymes, the fragment could be cloned into an appropriate expression and/or cloning vector which has been similarly digested (e.g., pSport1, among others). The skilled artisan would appreciate that other plasmids could be equally substituted, and may be desirable in certain circumstances. The digested fragment and vector are then ligated using a DNA ligase, and then used to transform competent *E.coli* cells using methods provided herein and/or otherwise known in the art.

[0327] The 5' primer sequence for amplifying any additional N-terminal deletion mutants may be determined by reference to the following formula:

(S+(X * 3)) to ((S+(X * 3))+25),

wherein 'S' is equal to the nucleotide position of the initiating start codon of the HGPRBMY8 gene (SEQ ID NO:1), and 'X' is equal to the most N-terminal amino acid of the intended N-terminal deletion mutant. The first term provides the start 5' nucleotide position of the 5' primer, while the second term provides the end 3' nucleotide position of the 5' primer corresponding to sense strand of SEQ ID NO:1. Once the corresponding nucleotide positions of the primer are determined, the final nucleotide sequence may be created by the addition of applicable restriction site sequences to the 5' end of the sequence, for example. As referenced herein, the addition of other sequences to the 5' primer may be desired in certain circumstances (e.g., kozac sequences, etc.).

[0328] The 3' primer sequence for amplifying any additional N-terminal deletion mutants may be determined by reference to the following formula:

(S+(X * 3)) to ((S+(X * 3))-25),

wherein 'S' is equal to the nucleotide position of the initiating start codon of the HGPRBMY8 gene (SEQ ID NO:1), and 'X' is equal to the most C-terminal amino acid of the intended N-terminal deletion mutant. The first term provides the start 5' nucleotide position of the 3' primer, while the second term provides the end 3' nucleotide position of the 3' primer corresponding to the anti-sense strand of SEQ ID NO:1. Once the corresponding nucleotide positions of the primer are determined, the final nucleotide sequence may be created by the addition of applicable restriction site sequences to the 5' end of the sequence, for example. As referenced herein, the

addition of other sequences to the 3' primer may be desired in certain circumstances (e.g., stop codon sequences, etc.). The skilled artisan would appreciate that modifications of the above nucleotide positions may be necessary for optimizing PCR amplification.

[0329] The same general formulas provided above may be used in identifying the 5' and 3' primer sequences for amplifying any C-terminal deletion mutant of the present invention. Moreover, the same general formulas provided above may be used in identifying the 5' and 3' primer sequences for amplifying any combination of N-terminal and C-terminal deletion mutant of the present invention. The skilled artisan would appreciate that modifications of the above nucleotide positions may be necessary for optimizing PCR amplification.

[0330] In preferred embodiments, the following N-terminal HGPRBMY8 deletion polypeptides are encompassed by the present invention: M1-P508, T2-P508, S3-P508, T4-P508, C5-P508, T6-P508, N7-P508, S8-P508, T9-P508, R10-P508, E11-P508, S12-P508, N13-P508, S14-P508, S15-P508, H16-P508, T17-P508, C18-P508, M19-P508, P20-P508, L21-P508, S22-P508, K23-P508, M24-P508, P25-P508, I26-P508, S27-P508, L28-P508, A29-P508, H30-P508, G31-P508, I32-P508, I33-P508, R34-P508, S35-P508, T36-P508, V37-P508, L38-P508, V39-P508, I40-P508, F41-P508, L42-P508, A43-P508, A44-P508, S45-P508, F46-P508, V47-P508, G48-P508, N49-P508, I50-P508, V51-P508, L52-P508, A53-P508, L54-P508, V55-P508, L56-P508, Q57-P508, R58-P508, K59-P508, P60-P508, Q61-P508, L62-P508, L63-P508, Q64-P508, V65-P508, T66-P508, N67-P508, R68-P508, F69-P508, I70-P508, F71-P508, N72-P508, L73-P508, L74-P508, V75-P508, T76-P508, D77-P508, L78-P508, L79-P508, Q80-P508, I81-P508, S82-P508, L83-P508, V84-P508, A85-P508, P86-P508, W87-P508, V88-P508, V89-P508, A90-P508, T91-P508, S92-P508, V93-P508, P94-P508, L95-P508, F96-P508, W97-P508, P98-P508, L99-P508, N100-P508, S101-P508, H102-P508, F103-P508, C104-P508, T105-P508, A106-P508, L107-P508, V108-P508, S109-P508, L110-P508, T111-P508, H112-P508, L113-P508, F114-P508, A115-P508, F116-P508, A117-P508, S118-P508, V119-P508, N120-P508, T121-P508, I122-P508, V123-P508, L124-P508, V125-P508, S126-P508, V127-P508, D128-P508, R129-P508, Y130-P508, L131-P508, S132-P508, I133-

P508, I134-P508, H135-P508, P136-P508, L137-P508, S138-P508, Y139-P508, P140-P508, S141-P508, K142-P508, M143-P508, T144-P508, Q145-P508, R146-P508, R147-P508, G148-P508, Y149-P508, L150-P508, L151-P508, L152-P508, Y153-P508, G154-P508, T155-P508, W156-P508, I157-P508, V158-P508, A159-P508, I160-P508, L161-P508, Q162-P508, S163-P508, T164-P508, P165-P508, P166-P508, L167-P508, Y168-P508, G169-P508, W170-P508, G171-P508, Q172-P508, A173-P508, A174-P508, F175-P508, D176-P508, E177-P508, R178-P508, N179-P508, A180-P508, L181-P508, C182-P508, S183-P508, M184-P508, I185-P508, W186-P508, G187-P508, A188-P508, S189-P508, P190-P508, S191-P508, Y192-P508, T193-P508, I194-P508, L195-P508, S196-P508, V197-P508, V198-P508, S199-P508, F200-P508, I201-P508, V202-P508, I203-P508, P204-P508, L205-P508, I206-P508, V207-P508, M208-P508, I209-P508, A210-P508, C211-P508, Y212-P508, S213-P508, V214-P508, V215-P508, F216-P508, C217-P508, A218-P508, A219-P508, R220-P508, R221-P508, Q222-P508, H223-P508, A224-P508, L225-P508, L226-P508, Y227-P508, N228-P508, V229-P508, K230-P508, R231-P508, H232-P508, S233-P508, L234-P508, E235-P508, V236-P508, R237-P508, V238-P508, K239-P508, D240-P508, C241-P508, V242-P508, E243-P508, N244-P508, E245-P508, D246-P508, E247-P508, E248-P508, G249-P508, A250-P508, E251-P508, K252-P508, K253-P508, E254-P508, E255-P508, F256-P508, Q257-P508, D258-P508, E259-P508, S260-P508, E261-P508, F262-P508, R263-P508, R264-P508, Q265-P508, H266-P508, E267-P508, G268-P508, E269-P508, V270-P508, K271-P508, A272-P508, K273-P508, E274-P508, G275-P508, R276-P508, M277-P508, E278-P508, A279-P508, K280-P508, D281-P508, G282-P508, S283-P508, L284-P508, K285-P508, A286-P508, K287-P508, E288-P508, G289-P508, S290-P508, T291-P508, G292-P508, T293-P508, S294-P508, E295-P508, S296-P508, S297-P508, V298-P508, E299-P508, A300-P508, R301-P508, G302-P508, S303-P508, E304-P508, E305-P508, V306-P508, R307-P508, E308-P508, S309-P508, S310-P508, T311-P508, V312-P508, A313-P508, S314-P508, D315-P508, G316-P508, S317-P508, M318-P508, E319-P508, G320-P508, K321-P508, E322-P508, G323-P508, S324-P508, T325-P508, K326-P508, V327-P508, E328-P508, E329-P508, N330-P508, S331-P508, M332-P508, K333-P508, A334-P508, D335-

P508, K336-P508, G337-P508, R338-P508, T339-P508, E340-P508, V341-P508, N342-P508, Q343-P508, C344-P508, S345-P508, I346-P508, D347-P508, L348-P508, G349-P508, E350-P508, D351-P508, D352-P508, M353-P508, E354-P508, F355-P508, G356-P508, E357-P508, D358-P508, D359-P508, I360-P508, N361-P508, F362-P508, S363-P508, E364-P508, D365-P508, D366-P508, V367-P508, E368-P508, A369-P508, V370-P508, N371-P508, I372-P508, P373-P508, E374-P508, S375-P508, L376-P508, P377-P508, P378-P508, S379-P508, R380-P508, R381-P508, N382-P508, S383-P508, N384-P508, S385-P508, N386-P508, P387-P508, P388-P508, L389-P508, P390-P508, R391-P508, C392-P508, Y393-P508, Q394-P508, C395-P508, K396-P508, A397-P508, A398-P508, K399-P508, V400-P508, I401-P508, F402-P508, I403-P508, I404-P508, I405-P508, F406-P508, S407-P508, Y408-P508, V409-P508, L410-P508, S411-P508, L412-P508, G413-P508, P414-P508, Y415-P508, C416-P508, F417-P508, L418-P508, A419-P508, V420-P508, L421-P508, A422-P508, V423-P508, W424-P508, V425-P508, D426-P508, V427-P508, E428-P508, T429-P508, Q430-P508, V431-P508, P432-P508, Q433-P508, W434-P508, V435-P508, I436-P508, T437-P508, I438-P508, I439-P508, I440-P508, W441-P508, L442-P508, F443-P508, F444-P508, L445-P508, Q446-P508, C447-P508, C448-P508, I449-P508, H450-P508, P451-P508, Y452-P508, V453-P508, Y454-P508, G455-P508, Y456-P508, M457-P508, H458-P508, K459-P508, T460-P508, I461-P508, K462-P508, K463-P508, E464-P508, I465-P508, Q466-P508, D467-P508, M468-P508, L469-P508, K470-P508, K471-P508, F472-P508, F473-P508, C474-P508, K475-P508, E476-P508, K477-P508, P478-P508, P479-P508, K480-P508, E481-P508, D482-P508, S483-P508, H484-P508, P485-P508, D486-P508, L487-P508, P488-P508, G489-P508, T490-P508, E491-P508, G492-P508, G493-P508, T494-P508, E495-P508, G496-P508, K497-P508, I498-P508, V499-P508, P500-P508, S501-P508, and/or Y502-P508 of SEQ ID NO:2. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these N-terminal HGPRBMY8 deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0331] In preferred embodiments, the following C-terminal HGPRBMY8 deletion polypeptides are encompassed by the present invention: M1-P508, M1-F507,

M1-T506, M1-A505, M1-S504, M1-D503, M1-Y502, M1-S501, M1-P500, M1-V499, M1-I498, M1-K497, M1-G496, M1-E495, M1-T494, M1-G493, M1-G492, M1-E491, M1-T490, M1-G489, M1-P488, M1-L487, M1-D486, M1-P485, M1-H484, M1-S483, M1-D482, M1-E481, M1-K480, M1-P479, M1-P478, M1-K477, M1-E476, M1-K475, M1-C474, M1-F473, M1-F472, M1-K471, M1-K470, M1-L469, M1-M468, M1-D467, M1-Q466, M1-I465, M1-E464, M1-K463, M1-K462, M1-I461, M1-T460, M1-K459, M1-H458, M1-M457, M1-Y456, M1-G455, M1-Y454, M1-V453, M1-Y452, M1-P451, M1-H450, M1-I449, M1-C448, M1-C447, M1-Q446, M1-L445, M1-F444, M1-F443, M1-L442, M1-W441, M1-I440, M1-I439, M1-I438, M1-T437, M1-I436, M1-V435, M1-W434, M1-Q433, M1-P432, M1-V431, M1-Q430, M1-T429, M1-E428, M1-V427, M1-D426, M1-V425, M1-W424, M1-V423, M1-A422, M1-L421, M1-V420, M1-A419, M1-L418, M1-F417, M1-C416, M1-Y415, M1-P414, M1-G413, M1-L412, M1-S411, M1-L410, M1-V409, M1-Y408, M1-S407, M1-F406, M1-I405, M1-I404, M1-I403, M1-F402, M1-I401, M1-V400, M1-K399, M1-A398, M1-A397, M1-K396, M1-C395, M1-Q394, M1-Y393, M1-C392, M1-R391, M1-P390, M1-L389, M1-P388, M1-P387, M1-N386, M1-S385, M1-N384, M1-S383, M1-N382, M1-R381, M1-R380, M1-S379, M1-P378, M1-P377, M1-L376, M1-S375, M1-E374, M1-P373, M1-I372, M1-N371, M1-V370, M1-A369, M1-E368, M1-V367, M1-D366, M1-D365, M1-E364, M1-S363, M1-F362, M1-N361, M1-I360, M1-D359, M1-D358, M1-E357, M1-G356, M1-F355, M1-E354, M1-M353, M1-D352, M1-D351, M1-E350, M1-G349, M1-L348, M1-D347, M1-I346, M1-S345, M1-C344, M1-Q343, M1-N342, M1-V341, M1-E340, M1-T339, M1-R338, M1-G337, M1-K336, M1-D335, M1-A334, M1-K333, M1-M332, M1-S331, M1-N330, M1-E329, M1-E328, M1-V327, M1-K326, M1-T325, M1-S324, M1-G323, M1-E322, M1-K321, M1-G320, M1-E319, M1-M318, M1-S317, M1-G316, M1-D315, M1-S314, M1-A313, M1-V312, M1-T311, M1-S310, M1-S309, M1-E308, M1-R307, M1-V306, M1-E305, M1-E304, M1-S303, M1-G302, M1-R301, M1-A300, M1-E299, M1-V298, M1-S297, M1-S296, M1-E295, M1-S294, M1-T293, M1-G292, M1-T291, M1-S290, M1-G289, M1-E288, M1-K287, M1-A286, M1-K285, M1-L284, M1-S283, M1-G282, M1-D281, M1-K280, M1-A279, M1-E278, M1-M277, M1-R276, M1-G275, M1-E274, M1-K273, M1-A272, M1-

K271, M1-V270, M1-E269, M1-G268, M1-E267, M1-H266, M1-Q265, M1-R264, M1-R263, M1-F262, M1-E261, M1-S260, M1-E259, M1-D258, M1-Q257, M1-F256, M1-E255, M1-E254, M1-K253, M1-K252, M1-E251, M1-A250, M1-G249, M1-E248, M1-E247, M1-D246, M1-E245, M1-N244, M1-E243, M1-V242, M1-C241, M1-D240, M1-K239, M1-V238, M1-R237, M1-V236, M1-E235, M1-L234, M1-S233, M1-H232, M1-R231, M1-K230, M1-V229, M1-N228, M1-Y227, M1-L226, M1-L225, M1-A224, M1-H223, M1-Q222, M1-R221, M1-R220, M1-A219, M1-A218, M1-C217, M1-F216, M1-V215, M1-V214, M1-S213, M1-Y212, M1-C211, M1-A210, M1-I209, M1-M208, M1-V207, M1-I206, M1-L205, M1-P204, M1-I203, M1-V202, M1-I201, M1-F200, M1-S199, M1-V198, M1-V197, M1-S196, M1-L195, M1-I194, M1-T193, M1-Y192, M1-S191, M1-P190, M1-S189, M1-A188, M1-G187, M1-W186, M1-I185, M1-M184, M1-S183, M1-C182, M1-L181, M1-A180, M1-N179, M1-R178, M1-E177, M1-D176, M1-F175, M1-A174, M1-A173, M1-Q172, M1-G171, M1-W170, M1-G169, M1-Y168, M1-L167, M1-P166, M1-P165, M1-T164, M1-S163, M1-Q162, M1-L161, M1-I160, M1-A159, M1-V158, M1-I157, M1-W156, M1-T155, M1-G154, M1-Y153, M1-L152, M1-L151, M1-L150, M1-Y149, M1-G148, M1-R147, M1-R146, M1-Q145, M1-T144, M1-M143, M1-K142, M1-S141, M1-P140, M1-Y139, M1-S138, M1-L137, M1-P136, M1-H135, M1-I134, M1-I133, M1-S132, M1-L131, M1-Y130, M1-R129, M1-D128, M1-V127, M1-S126, M1-V125, M1-L124, M1-V123, M1-I122, M1-T121, M1-N120, M1-V119, M1-S118, M1-A117, M1-F116, M1-A115, M1-F114, M1-L113, M1-H112, M1-T111, M1-L110, M1-S109, M1-V108, M1-L107, M1-A106, M1-T105, M1-C104, M1-F103, M1-H102, M1-S101, M1-N100, M1-L99, M1-P98, M1-W97, M1-F96, M1-L95, M1-P94, M1-V93, M1-S92, M1-T91, M1-A90, M1-V89, M1-V88, M1-W87, M1-P86, M1-A85, M1-V84, M1-L83, M1-S82, M1-I81, M1-Q80, M1-L79, M1-L78, M1-D77, M1-T76, M1-V75, M1-L74, M1-L73, M1-N72, M1-F71, M1-I70, M1-F69, M1-R68, M1-N67, M1-T66, M1-V65, M1-Q64, M1-L63, M1-L62, M1-Q61, M1-P60, M1-K59, M1-R58, M1-Q57, M1-L56, M1-V55, M1-L54, M1-A53, M1-L52, M1-V51, M1-I50, M1-N49, M1-G48, M1-V47, M1-F46, M1-S45, M1-A44, M1-A43, M1-L42, M1-F41, M1-I40, M1-V39, M1-L38, M1-V37, M1-T36, M1-S35, M1-R34, M1-I33, M1-I32, M1-G31, M1-H30, M1-A29, M1-L28, M1-S27, M1-I26, M1-P25, M1-M24,

M1-K23, M1-S22, M1-L21, M1-P20, M1-M19, M1-C18, M1-T17, M1-H16, M1-S15, M1-S14, M1-N13, M1-S12, M1-E11, M1-R10, M1-T9, M1-S8, and/or M1-N7 of SEQ ID NO:2. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these C-terminal HGPRBMY8 deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0332] Alternatively, preferred polypeptides of the present invention may comprise polypeptide sequences corresponding to, for example, internal regions of the HGPRBMY8 polypeptide (e.g., any combination of both N- and C- terminal HGPRBMY8 polypeptide deletions) of SEQ ID NO:2. For example, internal regions could be defined by the equation: amino acid NX to amino acid CX, wherein NX refers to any N-terminal deletion polypeptide amino acid of HGPRBMY8 (SEQ ID NO:2), and where CX refers to any C-terminal deletion polypeptide amino acid of HGPRBMY8 (SEQ ID NO:2). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of these polypeptides as an immunogenic and/or antigenic epitope as described elsewhere herein.

EXAMPLE 11

METHOD OF ENHANCING THE BIOLOGICAL ACTIVITY/FUNCTIONAL CHARACTERISTICS OF INVENTION THROUGH MOLECULAR EVOLUTION.

[0333] Although many of the most biologically active proteins known are highly effective for their specified function in an organism, they often possess characteristics that make them undesirable for transgenic, therapeutic, pharmaceutical, and/or industrial applications. Among these traits, a short physiological half-life is the most prominent problem, and is present either at the level of the protein, or the level of the proteins mRNA. The ability to extend the half-life, for example, would be particularly important for a proteins use in gene therapy, transgenic animal production, the bioprocess production and purification of the protein, and use of the protein as a chemical modulator among others. Therefore, there is a need to identify novel variants of isolated proteins possessing characteristics which enhance their application as a therapeutic for treating diseases of animal origin, in addition to the proteins applicability to common industrial and pharmaceutical applications.

[0334] Thus, one aspect of the present invention relates to the ability to enhance specific characteristics of invention through directed molecular evolution. Such an enhancement may, in a non-limiting example, benefit the inventions utility as an essential component in a kit, the inventions physical attributes such as its solubility, structure, or codon optimization, the inventions specific biological activity, including any associated enzymatic activity, the proteins enzyme kinetics, the proteins Ki, Kcat, Km, Vmax, Kd, protein-protein activity, protein-DNA binding activity, antagonist/inhibitory activity (including direct or indirect interaction), agonist activity (including direct or indirect interaction), the proteins antigenicity (e.g., where it would be desirable to either increase or decrease the antigenic potential of the protein), the immunogenicity of the protein, the ability of the protein to form dimers, trimers, or multimers with either itself or other proteins, the antigenic efficacy of the invention, including its subsequent use a preventative treatment for disease or disease states, or as an effector for targeting diseased genes. Moreover, the ability to enhance specific characteristics of a protein may also be applicable to changing the characterized activity of an enzyme to an activity completely unrelated to its initially characterized activity. Other desirable enhancements of the invention would be specific to each individual protein, and would thus be well known in the art and contemplated by the present invention.

[0335] For example, an engineered G-protein coupled receptor may be constitutively active upon binding of its cognate ligand. Alternatively, an engineered G-protein coupled receptor may be constitutively active in the absence of ligand binding. In yet another example, an engineered GPCR may be capable of being activated with less than all of the regulatory factors and/or conditions typically required for GPCR activation (e.g., ligand binding, phosphorylation, conformational changes, etc.). Such GPCRs would be useful in screens to identify GPCR modulators, among other uses described herein.

[0336] Directed evolution is comprised of several steps. The first step is to establish a library of variants for the gene or protein of interest. The most important step is to then select for those variants that entail the activity you wish to identify. The design of the screen is essential since your screen should be selective enough to

eliminate non-useful variants, but not so stringent as to eliminate all variants. The last step is then to repeat the above steps using the best variant from the previous screen. Each successive cycle, can then be tailored as necessary, such as increasing the stringency of the screen, for example.

[0337] Over the years, there have been a number of methods developed to introduce mutations into macromolecules. Some of these methods include, random mutagenesis, "error-prone" PCR, chemical mutagenesis, site-directed mutagenesis, and other methods well known in the art (for a comprehensive listing of current mutagenesis methods, see Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982)). Typically, such methods have been used, for example, as tools for identifying the core functional region(s) of a protein or the function of specific domains of a protein (if a multi-domain protein). However, such methods have more recently been applied to the identification of macromolecule variants with specific or enhanced characteristics.

[0338] Random mutagenesis has been the most widely recognized method to date. Typically, this has been carried out either through the use of "error-prone" PCR (as described in Moore, J., et al, Nature Biotechnology 14:458, (1996), or through the application of randomized synthetic oligonucleotides corresponding to specific regions of interest (as described by Derbyshire, K.M. et al, Gene, 46:145-152, (1986), and Hill, DE, et al, Methods Enzymol., 55:559-568, (1987). Both approaches have limits to the level of mutagenesis that can be obtained. However, either approach enables the investigator to effectively control the rate of mutagenesis. This is particularly important considering the fact that mutations beneficial to the activity of the enzyme are fairly rare. In fact, using too high a level of mutagenesis may counter or inhibit the desired benefit of a useful mutation.

[0339] While both of the aforementioned methods are effective for creating randomized pools of macromolecule variants, a third method, termed "DNA Shuffling", or "sexual PCR" (WPC, Stemmer, PNAS, 91:10747, (1994)) has recently been elucidated. DNA shuffling has also been referred to as "directed molecular evolution", "exon-shuffling", "directed enzyme evolution", "in vitro evolution", and "artificial evolution". Such reference terms are known in the art and are encompassed

by the invention. This new, preferred, method apparently overcomes the limitations of the previous methods in that it not only propagates positive traits, but simultaneously eliminates negative traits in the resulting progeny.

[0340] DNA shuffling accomplishes this task by combining the principal of in vitro recombination, along with the method of "error-prone" PCR. In effect, you begin with a randomly digested pool of small fragments of your gene, created by Dnase I digestion, and then introduce said random fragments into an "error-prone" PCR assembly reaction. During the PCR reaction, the randomly sized DNA fragments not only hybridize to their cognate strand, but also may hybridize to other DNA fragments corresponding to different regions of the polynucleotide of interest – regions not typically accessible via hybridization of the entire polynucleotide. Moreover, since the PCR assembly reaction utilizes "error-prone" PCR reaction conditions, random mutations are introduced during the DNA synthesis step of the PCR reaction for all of the fragments -further diversifying the potential hybridization sites during the annealing step of the reaction.

[0341] A variety of reaction conditions could be utilized to carry-out the DNA shuffling reaction. However, specific reaction conditions for DNA shuffling are provided, for example, in PNAS, 91:10747, (1994). Briefly, the DNA substrate to be subjected to the DNA shuffling reaction is prepared. Preparation may be in the form of simply purifying the DNA from contaminating cellular material, chemicals, buffers, oligonucleotide primers, deoxynucleotides, RNAs, etc., and may entail the use of DNA purification kits as those provided by Qiagen, Inc., or by the Promega, Corp., for example.

[0342] Once the DNA substrate has been purified, it would be subjected to Dnase I digestion. About 2-4µg of the DNA substrate(s) would be digested with 0.0015 units of Dnase I (Sigma) per microliter in 100µl of 50mM Tris-HCL, pH 7.4/1mM MgCl2 for 10-20 min. at room temperature. The resulting fragments of 10-50bp could then be purified by running them through a 2% low-melting point agarose gel by electrophoresis onto DE81 ion-exchange paper (Whatmann) or could be purified using Microcon concentrators (Amicon) of the appropriate molecular weight cutoff, or could use oligonucleotide purification columns (Qiagen), in addition to

other methods known in the art. If using DE81 ion-exchange paper, the 10-50bp fragments could be eluted from said paper using 1M NaCl, followed by ethanol precipitation.

[0343] The resulting purified fragments would then be subjected to a PCR assembly reaction by re-suspension in a PCR mixture containing: 2mM of each dNTP, 2.2mM MgCl2, 50 mM KCl, 10mM Tris•HCL, pH 9.0, and 0.1% Triton X-100, at a final fragment concentration of 10-30ng/µl. No primers are added at this point. *Taq* DNA polymerase (Promega) would be used at 2.5 units per 100ul of reaction mixture. A PCR program of 94 C for 60s; 94 C for 30s, 50-55 C for 30s, and 72 C for 30s using 30-45 cycles, followed by 72 C for 5min using an MJ Research (Cambridge, MA) PTC-150 thermocycler. After the assembly reaction is completed, a 1:40 dilution of the resulting primerless product would then be introduced into a PCR mixture (using the same buffer mixture used for the assembly reaction) containing 0.8um of each primer and subjecting this mixture to 15 cycles of PCR (using 94 C for 30s, 50 C for 30s, and 72 C for 30s). The referred primers would be primers corresponding to the nucleic acid sequences of the polynucleotide(s) utilized in the shuffling reaction. Said primers could consist of modified nucleic acid base pairs using methods known in the art and referred to else where herein, or could contain additional sequences (i.e., for adding restriction sites, mutating specific base-pairs, etc.).

[0344] The resulting shuffled, assembled, and amplified product can be purified using methods well known in the art (e.g., Qiagen PCR purification kits) and then subsequently cloned using appropriate restriction enzymes.

[0345] Although a number of variations of DNA shuffling have been published to date, such variations would be obvious to the skilled artisan and are encompassed by the invention. The DNA shuffling method can also be tailored to the desired level of mutagenesis using the methods described by Zhao, et al. (Nucl Acid Res., 25(6):1307-1308, (1997).

[0346] As described above, once the randomized pool has been created, it can then be subjected to a specific screen to identify the variant possessing the desired characteristic(s). Once the variant has been identified, DNA corresponding to the

variant could then be used as the DNA substrate for initiating another round of DNA shuffling. This cycle of shuffling, selecting the optimized variant of interest, and then re-shuffling, can be repeated until the ultimate variant is obtained. Examples of model screens applied to identify variants created using DNA shuffling technology may be found in the following publications: J. C., Moore, et al., J. Mol. Biol., 272:336-347, (1997), F.R., Cross, et al., Mol. Cell. Biol., 18:2923-2931, (1998), and A. Crameri., et al., Nat. Biotech., 15:436-438, (1997).

[0347] DNA shuffling has several advantages. First, it makes use of beneficial mutations. When combined with screening, DNA shuffling allows the discovery of the best mutational combinations and does not assume that the best combination contains all the mutations in a population. Secondly, recombination occurs simultaneously with point mutagenesis. An effect of forcing DNA polymerase to synthesize full-length genes from the small fragment DNA pool is a background mutagenesis rate. In combination with a stringent selection method, enzymatic activity has been evolved up to 16,000 fold increase over the wild-type form of the enzyme. In essence, the background mutagenesis yielded the genetic variability on which recombination acted to enhance the activity.

[0348] A third feature of recombination is that it can be used to remove deleterious mutations. As discussed above, during the process of the randomization, for every one beneficial mutation, there may be at least one or more neutral or inhibitory mutations. Such mutations can be removed by including in the assembly reaction an excess of the wild-type random-size fragments, in addition to the random-size fragments of the selected mutant from the previous selection. During the next selection, some of the most active variants of the polynucleotide/polypeptide/enzyme, should have lost the inhibitory mutations.

[0349] Finally, recombination enables parallel processing. This represents a significant advantage since there are likely multiple characteristics that would make a protein more desirable (e.g. solubility, activity, etc.). Since it is increasingly difficult to screen for more than one desirable trait at a time, other methods of molecular evolution tend to be inhibitory. However, using recombination, it would be possible to combine the randomized fragments of the best representative variants for the

various traits, and then select for multiple properties at once.

[0350] DNA shuffling can also be applied to the polynucleotides and polypeptides of the present invention to decrease their immunogenicity in a specified host. For example, a particular variant of the present invention may be created and isolated using DNA shuffling technology. Such a variant may have all of the desired characteristics, though may be highly immunogenic in a host due to its novel intrinsic structure. Specifically, the desired characteristic may cause the polypeptide to have a non-native structure which could no longer be recognized as a "self" molecule, but rather as a "foreign", and thus activate a host immune response directed against the novel variant. Such a limitation can be overcome, for example, by including a copy of the gene sequence for a xenobiotic ortholog of the native protein in with the gene sequence of the novel variant gene in one or more cycles of DNA shuffling. The molar ratio of the ortholog and novel variant DNAs could be varied accordingly. Ideally, the resulting hybrid variant identified would contain at least some of the coding sequence which enabled the xenobiotic protein to evade the host immune system, and additionally, the coding sequence of the original novel variant that provided the desired characteristics.

[0351] Likewise, the invention encompasses the application of DNA shuffling technology to the evolution of polynucleotides and polypeptides of the invention, wherein one or more cycles of DNA shuffling include, in addition to the gene template DNA, oligonucleotides coding for known allelic sequences, optimized codon sequences, known variant sequences, known polynucleotide polymorphism sequences, known ortholog sequences, known homologue sequences, additional homologous sequences, additional non-homologous sequences, sequences from another species, and any number and combination of the above.

[0352] In addition to the described methods above, there are a number of related methods that may also be applicable, or desirable in certain cases. Representative among these are the methods discussed in PCT applications WO 98/31700, and WO 98/32845, which are hereby incorporated by reference. Furthermore, related methods can also be applied to the polynucleotide sequences of the present invention in order to evolve invention for creating ideal variants for use in

gene therapy, protein engineering, evolution of whole cells containing the variant, or in the evolution of entire enzyme pathways containing polynucleotides of the invention as described in PCT applications WO 98/13485, WO 98/13487, WO 98/27230, WO 98/31837, and Crameri, A., et al., Nat. Biotech., 15:436-438, (1997), respectively.

[0353] Additional methods of applying "DNA Shuffling" technology to the polynucleotides and polypeptides of the present invention, including their proposed applications, may be found in US Patent No. 5,605,793; PCT Application No. WO 95/22625; PCT Application No. WO 97/20078; PCT Application No. WO 97/35966; and PCT Application No. WO 98/42832; PCT Application No. WO 00/09727 specifically provides methods for applying DNA shuffling to the identification of herbicide selective crops which could be applied to the polynucleotides and polypeptides of the present invention; additionally, PCT Application No. WO 00/12680 provides methods and compositions for generating, modifying, adapting, and optimizing polynucleotide sequences that confer detectable phenotypic properties on plant species; each of the above are hereby incorporated in their entirety herein for all purposes.

EXAMPLE 12

METHOD OF DISCOVERING ADDITIONAL SINGLE NUCLEOTIDE POLYMORPHISMS (SNPS) OF THE PRESENT INVENTION

[0354] Additional SNPs may be discovered in the polynucleotides of the present invention based on comparative DNA sequencing of PCR products derived from genomic DNA from multiple individuals. The genomic DNA samples may be purchased from Coriell Institute (Collingswood, NJ). PCR amplicons may be designed to cover the entire coding region of the exons using the Primer3 program (Rozen S 2000). Exon-intron structure of candidate genes and intron sequences may be obtained by blastn search of Genbank cDNA sequences against the human genome draft sequences. The sizes of these PCR amplicons will vary according to the exon-intron structure. All the samples may be amplified from genomic DNA (20 ng) in reactions (50 µl) containing 10 mM Tris-Cl pH 8.3, 50 mM KCl, 2.5 mM $MgCl_2$, 150

uM dNTPs, 3 uM PCR primers, and 3.75 U TaqGold DNA polymerase (PE Biosystems).

[0355] PCR is performed in MJ Research Tetrad machines under a cycling condition of 94 degrees 10 min, 30 cycles of 94 degrees 30 sec, 60 degrees 30sec, and 72 degrees 30 sec, followed by 72 degrees 7 min. PCR products may be purified using QIAquick PCR purification kit (Qiagen), and may be sequenced by the dye-terminator method using PRISM 3700 automated DNA sequencer (Applied Biosystems, Foster City, CA) following the manufacturer's instruction outlined in the Owner's Manual (which is hereby incorporated herein by reference in its entirety). Sequencing results may be analyzed for the presence of polymorphisms using PolyPhred software(Nickerson DA 1997; Rieder MJ 1999). All the sequence traces of potential polymorphisms may be visually inspected to confirm the presence of SNPs.

[0356] Alternative methods for identifying SNPs of the present invention are known in the art. One such method involves resequencing of target sequences from individuals of diverse ethnic and geographic backgrounds by hybridization to probes immobilized to microfabricated arrays. The strategy and principles for the design and use of such arrays are generally described in WO 95/11995.

[0357] A typical probe array used in such an analysis would have two groups of four sets of probes that respectively tile both strands of a reference sequence. A first probe set comprises a plurality of probes exhibiting perfect complementarily with one of the reference sequences. Each probe in the first probe set has an interrogation position that corresponds to a nucleotide in the reference sequence. That is, the interrogation position is aligned with the corresponding nucleotide in the reference sequence, when the probe and reference sequence are aligned to maximize complementarily between the two. For each probe in the first set, there are three corresponding probes from three additional probe sets. Thus, there are four probes corresponding to each nucleotide in the reference sequence. The probes from the three additional probe sets would be identical to the corresponding probe from the first probe set except at the interrogation position, which occurs in the same position in each of the four corresponding probes from the four probe sets, and is occupied by

a different nucleotide in the four probe sets. In the present analysis, probes may be nucleotides long. Arrays tiled for multiple different references sequences may be included on the same substrate.

[0358] Publicly available sequences for a given gene can be assembled into Gap4 (http:/ /www .biozentrum. unibas.ch/-biocomp/staden/Overview .html). PCR primers covering each exon, could be designed, for example, using Primer 3 (httP://www-genome.wi.mit.edu/cgi- bin/primer/primer3.cgi). Primers would not be designed in regions where there are sequence discrepancies between reads. Genomic DNA could be amplified from at least two individuals using 2.5 pmol each primer, 1.5 mM MgC12, 100 ~M dNTPs, 0.75 ~M AmpliTaq GOLD polymerase, and about 19ng DNA in a 15 ul reaction. Reactions could be assembled using a PACKARD MultiPROBE robotic pipetting station and then put in MJ 96-well tetrad thermocyclers (96°C for minutes, followed by cycles of 96°C for seconds, 59°C for 2 minutes, and 72°C for 2 minutes). A subset of the PCR assays for each individual could then be run on 3% NuSieve gels in 0.5X TBE to confirm that the reaction worked.

[0359] For a given DNA, 5ul (about 50 ng) of each PCR or RT -PCR product could be pooled (Final volume = 150-200 ul). The products can be purified using QiaQuick PCR purification from Qiagen. The samples would then be eluted once in 35ul sterile water and 4 ul lOX One-Phor-All buffer (Pharmacia). The pooled samples are then digested with 0.2u DNaseI (Promega) for 10 minutes at 37°C and then labeled with 0.5 nmols biotin-N6- ddATP and 15u Terminal Transferase (GibcoBRL Life Technology) for 60 minutes at 37°C. Both fragmentation and labeling reactions could be terminated by incubating the pooled sample for 15 minutes at 100°C.

[0360] Low-density DNA chips (Affymetrix,CA) may be hybridized following the manufacturer's instructions. Briefly, the hybridization cocktail consisted of 3M TMACI, mM Tris pH 7.8, 0.01% Triton X-100, 100 mg/ml herring sperm DNA {Gibco BRL), 200 pM control biotin-labeled oligo. The processed PCR products are then denatured for 7 minutes at 100°C and then added to prewarmed {37°C) hybridization solution. The chips are hybridized overnight at 44°C. Chips are

ished in 1X SSPET and 6X SSPET followed by staining with 2 ug/ml SARPE and 0.5 mg/ml acetylated BSA in 200 ul of 6X SSPET for 8 minutes at room temperature. Chips are scanned using a Molecular Dynamics scanner.

[0361] Chip image files may be analyzed using Ulysses {Affymetrix, CA) which uses four algorithms to identify potential polymorphisms. Candidate polymorphisms may be visually inspected and assigned a confidence value: where high confidence candidates display all three genotypes, while likely candidates show only two genotypes {homozygous for reference sequence and heterozygous for reference and variant). Some of the candidate polymorphisms may be confirmed by ABI sequencing. Identified polymorphisms could then be compared to several databases to determine if they are novel.

EXAMPLE 13

METHOD OF DETERMINING THE ALLELE FREQUENCY FOR EACH SNP OF THE PRESENT INVENTION.

[0362] Allele frequencies of these polymorphisms may be determined by genotyping various DNA samples (Coriell Institute; Collingswood, NJ) using FP-TDI assay (Chen X 1999). Automated genotyping calls may be made with an allele calling software developed by Joel Hirschorn (Whitehead Institute/MIT Center for Genome Research, personal communication).

[0363] Briefly, the no template controls (NTCs) may be labeled accordingly in column C. The appropriate cells may be completed in column L indicating whether REF (homozygous ROX) or VAR (homozygous TAMRA) are expected to be rare genotypes (<10% of all samples) – the latter is important in helping the program to identify rare homozygotes. The number of 96 well plates genotyped in cell P2 are noted (generally between 0.5 and 4) - the program works best if this is accurate. No more than 384 samples can be analyzed at a time. The pairs of mP values from the LJL may be pasted into columns E and F; making sure there may be no residual data is left at the bottom fewer than 384 data points are provided. The DNA names may be provided in columns A, B or C; column I will be a concatenation of columns A, B and C. In addition, the well numbers for each sample may be also provided in column D.

[0364] With the above information provided, the program should

automatically cluster the points and identify genotypes. The program works by converting the mP values into polar coordinates (distance from origin and angle from origin) with the angle being on a scale from 0 to 2; heterozygotes are placed as close to 1 as possible.

[0365] The cutoff values in columns L and M may be adjusted as desired.

[0366] Expert parameters: The most important parameters are the maximum angle for REF and minimum angle for VAR. These parameters may need to be changed in a particularly skewed assay which may be observed when an REF or VAR cluster is close to an angle of 1 and has called as a failed or HETs.

[0367] Other parameters are low and high cutoffs that are used to determine which points are considered for the determination of edges of the clusters. With small numbers of data points, the high cutoff may need to be increased (to 500 or so). This may be the right thing to do for every assay, but certainly when the program fails to identify a small cluster with high signal.

[0368] NTC TAMRA and ROX indicate the position of the no template control or failed samples as estimated by the computer algorithm.

[0369] No signal = mP< is the threshold below which points are automatically considered failures. "Throw out points with signal above" is the TAMRA or ROX mP value above which points are considered failures. The latter may occasionally need to be adjusted from 250 to 300, but caveat emptor for assays with signals >250. 'Lump' or 'split' describes a subtle difference in the way points are grouped into clusters. Lump generally is better. 'HETs expected' in the rare case where only homozygotes of either class are expected (e.g. a study of X chromosome SNPs in males), change this to "N".

[0370] Notes on method of clustering: The origin is defined by the NTCs or other low signal points (the position of the origin is shown as "NTC TAMRA" and "NTC ROX"); the points with very low or high signal are not considered initially. The program finds the point farthest from the origin and calls that a HET; the ROX/TAMRA ratio is calculated from this point, placing the heterozygotes at 45 degrees from the origin (an angle of "1"). The angles from the origin are calculated (the scale ranges from 0 to 2) and used to define clusters. A histogram of angles is

generated. The cluster boundaries are defined by an algorithm that takes into account the shape of the histogram. The homozygote clusters are defined as the leftmost and rightmost big clusters (unless the allele is specified as being rare, in which case the cluster need not be big). The heterozygote is the biggest cluster in between the REF and VAR. If there are two equal clusters, the one best-separated from REF and VAR is called HET. All other clusters are failed. Some fine tuning is applied to lump in scattered points on the edges of the clusters (if "Lump" is selected). The boundaries of the clusters are "Angles" in column L.

[0371] Once the clusters are defined, the interquartile distance of signal intensity is defined for each cluster. Points falling more than 3 or 4 interquartiles from the mean are excluded. (These are the "Signal cutoffs" in column M).

[0372] Allele frequency of the B1 receptor R317Q variant (AE103s1) is as follows. 7% in African Americans (7/94), 0% in Caucasians (0/94), 0% in Asians (0/60), and 0% in Amerindians (0/20). Higher frequency of this form in African Americans than in Caucasians matches the profile of a potential genetic risk factor for angioedema, which is observed more frequently in African Americans than in Caucasians (Brown NJ 1996; Brown NJ 1998; Agostoni A 1999; Coats 2000).

[0373] The invention encompasses additional methods of determinig the allelic frequency of the SNPs of the present invention. Such methods may be known in the art, some of which are described elsewhere herein.

EXAMPLE 14

ALTERNATIVE METHODS OF DETECTING POLYMORPHISMS ENCOMPASSED BY THE PRESENT INVENTION.

Preparation of Samples

[0374] Polymorphisms are detected in a target nucleic acid from an individual being analyzed. For assay of genomic DNA, virtually any biological sample (other than pure red blood cells) is suitable. For example, convenient tissue samples include whole blood, semen, saliva, tears, urine, fecal material, sweat, buccal, skin and hair. For assay of cDNA or mRNA, the tissue sample must be obtained from an organ in which the target nucleic acid is expressed. For example, if the target nucleic acid is a cytochrome P450, the liver is a suitable source.

[0375] Many of the methods described below require amplification of DNA from target samples. This can be accomplished by e.g., PCR. See generally PCR Technology: Principles and Applications for DNA Amplification (ed. H.A. Erlich, Freeman Press, NY, NY, 1992); PCR Protocols: A Guide to Methods and Applications (eds. Innis, et al., Academic Press, San Diego, CA, 1990); Mattila et al., Nucleic Acids Res. 19, 4967 (1991); Eckert et al., PCR Methods and Applications 1, (1991); PCR (eds. McPherson et al., IRL Press, Oxford); and U.S. Patent 4,683,202.

[0376] Other suitable amplification methods include the ligase chain reaction (LCR) (see Wu and Wallace, Genomics 4:560 (1989), Landegren et al., Science 241:1077 (1988), transcription amplification (Kwoh et al., Proc. Natl. Acad. Sci. USA 86, 1173 (1989), and self-sustained sequence replication (Guatelli et al., Proc. Nat. Acad. Sci. USA, 87:1874 (1990)) and nucleic acid based sequence amplification (NASBA). The latter two amplification methods involve isothermal reactions based on isothermal transcription, which produce both single stranded RNA (ssRNA) and double stranded DNA (dsDNA) as the amplification products in a ratio of about 30 or 100 to 1, respectively.

[0377] Additional methods of amplification are known in the art or are described elsewhere herein.

Detection of Polymorphisms in Target DNA

[0378] There are two distinct types of analysis of target DNA for detecting polymorphisms. The first type of analysis, sometimes referred to as de novo characterization, is carried out to identify polymorphic sites not previously characterized (i.e., to identify new polymorphisms). This analysis compares target sequences in different individuals to identify points ofvariation, i.e., polymorphic sites. By analyzing groups of individuals representing the greatest ethnic diversity among humans and greatest breed and species variety in plants and animals, patterns characteristic of the most common alleles/haplotypes of the locus can be identified, and the frequencies of such alleles/haplotypes in the population can be determined. Additional allelic frequencies can be determined for subpopulations characterized by criteria such as geography, race, or gender. The de novo identification ofpolymorphisms of the invention is described in the Examples section.

[0379] The second type of analysis determines which form(s) ofa characterized (known) polymorphism are present in individuals under test. Additional methods of analysis are known in the art or are described elsewhere herein.

Allele-Specific Probes

[0380] The design and use of allele-specific probes for analyzing polymorphisms is described by e.g., Saiki et al., Nature 324,163-166 (1986); Dattagupta, EP 235,726, Saiki, WO 89/11548. Allele-specific probes can be designed that hybridize to a segment of target DNA from one individual but do not hybridize to the corresponding segment from another individual due to the presence of different polymorphic forms in the respective segments from the two individuals. Hybridization conditions should be sufficiently stringent that there is a significant difference in hybridization intensity between alleles, and preferably an essentially binary response, whereby a probe hybridizes to only one of the alleles. Some probes are designed to hybridize to a segment of target DNA such that the polymorphic site aligns with a central position (e.g., in a 15-mer at the 7 position; in a 16-mer, at either the 8 or 9 position) of the probe. This design of probe achieves good discrimination in hybridization between different allelic forms.

[0381] A1lele-specific probes are often used in pairs, one member of a pair showing a perfect match to a reference form of a target sequence and the other member showing a perfect match to a variant form. Several pairs of probes can then be immobilized on the same support for simultaneous analysis of multiple polymorphisms within the same target sequence.

Tiling Arrays

[0382] The polymorphisms can also be identified by hybridization to nucleic acid arrays, some examples of which are described in WO 95/11995. The same arrays or different arrays can be used for analysis of characterized polymorphisms. WO 95/11995 also describes sub arrays that are optimized for detection of a variant form of a precharacterized polymorphism. Such a subarray contains probes designed to be complementary to a second reference sequence, which is an allelic variant of the first reference sequence. The second group of probes is designed by the same principles as described, except that the probes exhibit complementarity to the second reference

sequence. The inclusion of a second group (or further groups) can be particularly useful for analyzing short subsequences of the primary reference sequence in which multiple mutations are expected to occur within a short distance commensurate with the length of the probes (e.g., two or more mutations within 9 to bases).

Allele-Specific Primers

[0383] An allele-specific primer hybridizes to a site on target DNA overlapping a polymorphism and only primes amplification of an allelic form to which the primer exhibits perfect complementarity. See Gibbs, Nucleic Acid Res. 17,2427-2448 (1989). This primer is used in conjunction with a second primer which hybridizes at a distal site. Amplification proceeds from the two primers, resulting in a detectable product which indicates the particular allelic form is present. A control is usually performed with a second pair ofprimers, one ofwhich shows a single base mismatch at the polymorphic site and the other of which exhibits perfect complementarity to a distal site. The single-base mismatch prevents amplification and no detectable product is formed. The method works best when the mismatch is included in the 3'-most position of the oligonucleotide aligned with the polymorphism because this position is most destabilizing elongation from the primer (see, e.g., WO 93/22456).

Direct-Sequencing

[0384] The direct analysis of the sequence of polymorphisms of the present invention can be accomplished using either the dideoxy chain termination method or the Maxam - Gilbert method (see Sambrook et al., Molecular Cloning, A Laboratory Manual (2nd Ed., CSHP, New York 1989); Zyskind et al., Recombinant DNA Laboratory Manual, (Acad. Press, 1988)).

Denaturing Gradient Gel Electrophoresis

[0385] Amplification products generated using the polymerase chain reaction can be analyzed by the use of denaturing gradient gel electrophoresis. Different alleles can be identified based on the different sequence-dependent melting properties and electrophoretic migration of DNA in solution. Erlich, ed., PCR Technology. Principles and Applications for DNA Amplification, (W .H. Freeman and Co, New York, 1992), Chapter 7.

Single-Strand Conformation Polymorphism Analysis

[0386] Alleles of target sequences can be differentiated using single-strand conformation polymorphism analysis, which identifies base differences by alteration in electrophoretic migration of single stranded PCR products, as described in Orita et al., Proc. Nat. Acad. Sci. 86,2766-2770 (1989). Amplified PCR products can be generated as described above, and heated or otherwise denatured, to form single stranded amplification products. Single-stranded nucleic acids may refold or form secondary structures which are partially dependent on the base sequence. The different electrophoretic mobilities of single-stranded amplification products can be related to base-sequence differences between alleles of target sequences.

Single Base Extension

[0387] An alternative method for identifying and analyzing polymorphisms is based on single-base extension (SBE) of a fluorescently-labeled primer coupled with fluorescence resonance energy transfer (FRET) between the label of the added base and the label of the primer. Typically, the method, such as that described by Chen et al., (PNAS 94:10756-61 (1997), uses a locus-specific oligonucleotide primer labeled on the 5' terminus with 5-carboxyfluorescein (F AM). This labeled primer is designed so that the 3' end is immediately adjacent to the polymorphic site of interest. The labeled primer is hybridized to the locus, and single base extension of the labeled primer is performed with fluorescently-labeled dideoxyribonucleotides (ddNTPs) in dye-terminator sequencing fashion. An increase in fluorescence of the added ddNTP in response to excitation at the wavelength of the labeled primer is used to infer the identity of the added nucleotide.

[0388] The contents of all patents, patent applications, published PCT applications and articles, books, references, reference manuals and abstracts cited herein are hereby incorporated by reference in their entirety to more fully describe the state of the art to which the invention pertains.

[0389] As various changes can be made in the above-described subject matter without departing from the scope and spirit of the present invention, it is intended that all subject matter contained in the above description, or defined in the appended

claims, be interpreted as descriptive and illustrative of the present invention. Many modifications and variations of the present invention are possible in light of the above teachings.

REFERENCES

1. Rees, S., Brown, S., Stables, J.: "Reporter gene systems for the study of G Protein Coupled Receptor signalling in mammalian cells". In Milligan G. (ed.): Signal Transduction: A practical approach. Oxford: Oxford University Press, 1999: 171-221.
2. Alam, J., Cook, J.L.: "Reporter Genes: Application to the study of mammalian gene transcription". Anal. Biochem. 1990; 188: 245-254.
3. Selbie, L.A. and Hill, S.J.: "G protein-coupled receptor cross-talk: The fine-tuning of multiple receptor-signaling pathways". TiPs. 1998; 19: 87-93.
4. Boss, V., Talpade, D.J., and Murphy, T.J.: "Induction of NFAT mediated transcription by Gq-coupled Receptors in lympoid and non-lymphoid cells". JBC. 1996; 271: 10429-10432.
5. George, S.E., Bungay, B.J., and Naylor, L.H.: "Functional coupling of endogenous serotonin (5-HT1B) and calcitonin (C1a) receptors in CHO cells to a cyclic AMP-responsive luciferase reporter gene". J. Neurochem. 1997; 69: 1278-1285.
6. Suto, CM, Igna DM: "Selection of an optimal reporter for cell-based high throughput screening assays". J. Biomol. Screening. 1997; 2: 7-12.
7. Zlokarnik, G., Negulescu, P.A., Knapp, T.E., More, L., Burres, N., Feng, L., Whitney, M., Roemer, K., and Tsien, R. Y. "Quantitation of transcription and clonal selection of single living cells with a B-Lactamase Reporter". Science. 1998; 279: 84-88.
8. S. Fiering et. al., Genes Dev. 4, 1823 (1990).
9. J. Karttunen and N. Shastri, PNAS 88, 3972 (1991).
10. Hawes, B. E., Luttrell. L.M., van Biesen, T., and Lefkowitz, R.J. (1996) JBC 271, 12133-12136.
11. Gilman, A.G. (1987) Annul. Rev. Biochem. 56, 615-649.
12. Maniatis et al., Cold Spring Harbor Press, 1989.
13. Salcedo, R., Ponce, M.L., Young, H.A., Wasserman, K., Ward, J.M., Kleinman, H.K., Oppenheim, J.J., Murphy, W.J. "Human endothelial cells express CCF2 and respond to MCP-1: direct role of MCP-1 in angiogenesis and tumor progression". Blood. 2000; 96 (1): 34-40.
14. Sica, A., Saccani, A., Bottazzi, B., Bernasconi, S., Allavena, P., Gaetano, B., LaRossa, G., Scotton, C., Balkwill F., Mantovani, A. "Defective expression of the monocyte chemotactic protein 1 receptor CCF2 in macrophages associated with human ovarian carcinoma". J. Immunology. 2000; 164: 733-8.
15. Kypson, A., Hendrickson, S., Akhter, S., Wilson, K., McDonald, P., Lilly, R., Dolber, P., Glower, D., Lefkowitz, R., Koch, W. "Adenovirus-mediated gene transfer of the B2 AR to donor hearts enhances cardiac function". Gene Therapy. 1999; 6: 1298-304.
16. Dorn, G.W., Tepe, N.M., Lorenz, J.N., Kock, W.J., Ligget, S.B. "Low and high level transgenic expression of B2AR differentially affect cardiac hypertrophy and function in Galpha q-overexpressing mice". PNAS. 1999; 96: 6400-5.
17. J. Wess. "G protein coupled receptor: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition". FASEB. 1997; 11:346-354.

18. Whitney, M, Rockenstein, E, Cantin, G., Knapp, T., Zlokarnik, G., Sanders, P., Durick, K., Craig, F.F., and Negulescu, P.A. "A genome-wide functional assay of signal transduction in living mammalian cells". Nature Biotech. 1998; 16: 1329-1333.

19. BD Biosciences: FACS Vantage SE Training Manual. Part Number 11-11020-00 Rev. A. August 1999.

20. Chen, G., Jaywickreme, C., Way, J., Armour S., Queen K., Watson., C., Ignar, D., Chen, W.J., Kenakin, T. "Constitutive Receptor systems for drug discovery". J. Pharmacol. Toxicol. Methods 1999; 42: 199-206.

21. Blahos, J., Fischer,T., Brabet, I., Stauffer,D., Rovelli, G., Bockaert, J., and Pin, J.-P. "A novel Site on the G alpha-protein that Rocognized Heptahelical Receptors". J.Biol. Chem. 2001; 275, No. 5, 3262-69.

22. Offermanns, S. & Simon, M.I. "G alpha 15 and G alpha 16 Couple a Wide Variety of Receptors to Phospholipase C". J. Biol. Chem. 1995; 270, No. 25, 15175-80.